Global HIV Strategic Information Working Group

Biobehavioural Survey Guidelines

For Populations At Risk For HIV

UNAIDS

fhi360 THE SCIENCE OF IMPROVING LIVES

World Health Organization

CDC Centers for Disease Control and Prevention

PEPFAR

© World Health Organization 2017

Some rights reserved. This work is available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; https://creativecommons.org/licenses/by-nc-sa/3.0/igo).

The mark "CDC" is owned by the US Dept. of Health and Human Services and is used with permission. Use of this logo is not an endorsement by HHS or CDC of any particular product, service, or enterprise.

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition".

Any mediation relating to disputes arising under the licence shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization.

Suggested citation
WHO, CDC, UNAIDS, FHI 360. Biobehavioral survey guidelines for Populations at Risk for HIV. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data
CIP data are available at http://apps.who.int/iris.

Sales, rights and licensing
To purchase WHO publications, see http://apps.who.int/bookorders. To submit requests for commercial use and queries on rights and licensing, see http://www.who.int/about/licensing.

Third-party materials
If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

General disclaimers
The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by WHO to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall WHO be liable for damages arising from its use.

Printed in Switzerland

ISBN: 978-92-4-151301-2

Authorship and acknowledgements

Authors
Abu Abdul-Quader, Mark Berry, Trista Bingham, Janet Burnett, Maxia Dong, Amy Drake, Avi Hakim, Wolfgang Hladik, Angele Marandet, Anne McIntyre, Chris Murrill, Joyce Neal and Nita Patel of the US Centers for Disease Control and Prevention (CDC); Rajatashuvra Adhikary (formerly of FHI 360); Tobi Saidel of Partnership for Epidemic Analysis (PEMA) Partners; and Angela Kelly-Hanku of the University of New South Wales and of the Papua New Guinea Institute of Medical Research.

Editors
Abu Abdul-Quader, Mark Berry, Trista Bingham, Janet Burnett, Dana Dolan, Maxia Dong, Amy Drake, Avi Hakim, Wolfgang Hladik, Angele Marandet, Anne McIntyre, Chris Murrill, Joyce Neal and Nita Patel of CDC; Rajatashuvra Adhikary (formerly of FHI 360), Johannes van Dam and Steve Mills of FHI 360; staff of the Joint United Nations Programme on HIV/AIDS (UNAIDS); Jesus Garcia Calleja of WHO; Thomas Rehle of the Human Sciences Research Council (HSRC); Tobi Saidel of PEMA Partners; and Ted Alcorn of the Bill & Melinda Gates Foundation.

Reviewers
Maxia Dong, Shahul Ebrahim, Avi Hakim, Wolfgang Hladik, Amy Herman-Roloff, Andrea Kim, Rachel Kwezi, Sheryl Lyss, John Macom, Chris Murrill, Patrick Nadol, Sanny Chen Northbrook, Bharat Parekh, Nita Patel, Dimitri Prybylski, Ray Shiraishi and Peter Young of CDC; Rajatashuvra Adhikary (formerly of FHI 360), Timothy Mastro, Mike Merrigan, Steve Mills and Johannes van Dam of FHI 360; staff of UNAIDS; Jesus Garcia Calleja of WHO; Thomas Rehle: HSRC and University of Cape Town; Tobi Saidel of PEMA Partners; and Angela Kelly-Hanku of the Kirby Institute, University of New South Wales, and Sexual and Reproductive Health Unit, Papua New Guinea Institute of Medical Research.

We would like to thank the individuals who contributed to this document:
Ashley Burson and Laura Porter of CDC; Helen Coelho, Amanda Geller, Seseni Nu, Betty Treschitta and Almeta West of ICF International, Vanessa Brown of the Office of the US Global AIDS Coordinator; Maria Au, and Tisha Wheeler of the United States Agency for International Development (USAID), and Emily Crawford (formerly with USAID)

The findings, conclusions and recommendations in this report do not necessarily represent the official position of the CDC.

Funding to support this work is from the United States President's Emergency Plan for AIDS Relief (PEPFAR) and the Bill & Melinda Gates Foundation.

Foreword

To address a public health problem, you first have to measure it accurately. Biobehavioural surveys have proven to be invaluable tools for measuring and addressing HIV, which remains the world's biggest public health challenge. This current iteration of the *Biobehavioural survey guidelines* is a welcome addition to the list of useful documents targeting those who plan to conduct biobehavioural surveys of HIV and HIV-risk behaviours in their countries. The guidelines can be applied across different countries, and to this end the document provides questionnaire modules that can be adapted to various contexts. The guidelines are presented in a logical and coherent manner, covering all survey aspects, from conceptualization of the survey to dissemination of the report and data use.

The major focus is on key populations, which are often hidden and difficult to measure as part of general population-based surveys. This is particularly important because key populations are at high risk for HIV, and for exclusion from HIV and other health services. Estimating the size of these populations and their burden of HIV disease is extremely challenging, and these guidelines are a valuable resource for survey specialists as they undertake the surveys.

The guidelines fill a gap in providing tools for surveying HIV prevalence in key populations, and the included questionnaires may also inform general population surveys. The 2000 Behavioural Surveillance Survey guidelines, while still useful, needed to be updated with newer survey methodology techniques and to incorporate biomarker testing. The guidelines will also serve as a textbook for students interested in working for research institutions that embark on epidemiological surveys.

Currently, many researchers undertake surveys using country-specific indicators. These guidelines standardize the conduct of biobehavioural surveys to permit comparisons between as well as within countries over time. The use of common indicators allows for uniformity in the measurement of items and production of data that can be used by various global, regional, national and local actors in planning prevention and treatment services, tracking progress in the provision of HIV prevention and treatment services, and identifying gaps in access to services. The appendix on indicators will help scientists and data specialists to harmonize data management with a view to collaborating across countries using common yardsticks.

The authors of these guidelines considered key aspects of surveys, from survey planning, design, data collection, analysis, presentation of results and dissemination of reports to data use. These guidelines are a must-have for anyone planning to conduct surveillance, whether experienced or not. It is our hope that the guidelines will help to refine measurement of HIV and help countries to address the unmet needs of their communities, to further reduce the toll of the epidemic.

Olive Shisana
Hon Professor, University of Cape Town
President and CEO, Evidence Based Solutions

and

Chris Beyrer
Desmond M. Tutu Professor of Public Health and Human Rights,
Johns Hopkins University,
Bloomberg School of Public Health

Preface

Biobehavioural surveys (BBS) provide specific population-level estimates for the burden of HIV disease and HIV-related risk factors, and estimates for the coverage of prevention and treatment services for populations at increased risk for HIV. These key populations include men who have sex with men, sex workers, people who inject drugs, transgender individuals, prisoners and other vulnerable populations at increased risk for HIV infection. For many of these stigmatized and socially marginalized populations, there are no conventional sampling frames, meaning that complex sampling designs are needed for these populations. The most frequently used survey guidelines and tools to date are the *Behavioral surveillance surveys*, issued in 2000. However, new HIV prevention, care and treatment policies – coupled with the emergence of new data needs, methods and technologies – warranted a thorough update of the 2000 publication. Thus, the US Centers for Disease Control and Prevention, FHI 360, the Joint United Nations Programme on HIV/AIDS (UNAIDS) and the World Health Organization (WHO) are publishing these new *Biobehavioural survey guidelines for populations at risk for HIV*. This revised publication outlines the latest approaches and methodologies, and includes updated questionnaires for planning and conducting BBS.

The new BBS guidelines are a comprehensive resource that covers all survey aspects, from conceptualization, planning, undertaking and costing of the survey to dissemination of a report and use of data. The ultimate goal of these guidelines is to facilitate the collection of high-quality survey data for informed public health action.

The new guidelines:

- update the overall approach and methodology of BBS in light of advances made during the past two decades;
- improve the quality of BBS by providing comprehensive guidance, particularly for sampling methods and the collection of interview and biomarker data;
- increase the relevance of survey data for public health programming by ensuring the collection of representative and actionable data; and
- promote the use of survey findings to improve service delivery, monitoring and evaluation, and policy development.

This document includes several new topics and features:

- formative assessment covers the initial collection of information about a population, to inform how best to prepare and conduct a BBS.
- respondent driven sampling covers this peer-driven chain-referral sampling method, which is particularly useful for hard-to-sample populations, and is currently viewed as the most suitable probability-based sampling design.
- biomarker considerations covers the entire range of biological measurements, from HIV serology to viral load, HIV recency, and biomarkers of other sexually transmitted infections. The document emphasizes the potential of population-level, aggregate viral-load metrics, such as prevalence of unsuppressed viral load.
- population size estimation provides guidance on using integrated methods to estimate the number of members of a population.
- questionnaire modules are included for collecting data on a wide range of topics, including exposure to and uptake of HIV-related services. Accompanying this publication will be electronic ready-to-use questionnaires, with the aim of keeping the questionnaires up-to-date as standards and indicators change.
- indicators appendix lists standard and newly proposed indicators for both data and biomarker-related metrics.

These guidelines, built on the lessons learned by experts and implementers from around the world, are intended to serve as a one-stop resource for survey planning and implementation, and dissemination of findings. They are expected to improve the quality of survey data through better survey design and implementation, and to promote the standardization of data measures to improve the comparability of survey data. Most of all, the guidelines should make it easier to plan and conduct BBS. By providing sample documents and guidance on every step of the process, we hope that these guidelines strengthen the capacity of public health workers to collect the information they need in a timely manner, allowing for an effective and strategic response to stop the HIV epidemic among key populations.

Dr Shannon Hader
Director, Division of Global HIV and TB, Center for Global Health/CDC

Dr Gottfried Hirnschall
Director, Department of HIV/AIDS and Global Hepatitis Programme, WHO

Dr Luiz Loures
Deputy Executive Director, Programme Branch, UNAIDS.

Dr Timothy Mastro
Chief Science Officer, FHI 360

Abbreviations and Acronyms

ACASI	audio computer-assisted self-interview
ART	antiretroviral therapy
ARV	antiretroviral
BBS	biobehavioural survey
CAB	community advisory board
CAPI	computer-assisted personal interview
CASI	computer-assisted self-interview
CCS	conventional cluster sampling
CDC	Centers for Disease Control and Prevention
CI	confidence interval
CRC	capture-recapture method
CT	*Chlamydia trachomatis*
DBS	dried blood spot
DEFF	design effect
DFA	direct fluorescent antibody
DNA	deoxyribonucleic acid
EIA	enzyme immunoassay
EMoS	estimated measure of size
EPS	equal probability sampling
FSW	female sex worker
FP	family planning
GoC	game of contacts
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HPV	human papillomavirus
HSRC	Human Sciences Research Council
HSV-2	herpes simplex virus-2
ID	identification number
IEC	information, education and communication
IRB	institutional review board
MoS	measure of size
MSM	men who have sex with men
NAAT	nucleic acid amplification test
NG	*Neisseria gonorrhoeae*
NGO	nongovernmental organization
PCR	polymerase chain reaction
PEP	post-exposure prophylaxis
PMTCT	prevention of mother-to-child transmission
PPS	probability proportional to size
PrEP	pre-exposure prophylaxis
PSU	primary sampling unit
PWID	people who inject drugs
QA	quality assurance
QDS	Questionnaire Development System
REC	research ethics committee
RDS	respondent-driven sampling
RNA	ribonucleic acid
RPR	rapid plasma reagin
RS	random start
RT	rapid test
SOP	standard operating procedure
SI	sampling interval
SRS	simple random sampling
STD/STI	sexually transmitted disease/sexually transmitted infection
SW	sex worker
TB	tuberculosis
TG	transgender person
TLS	time-location sampling
UN	United Nations
UNAIDS	Joint United Nations Programme on HIV/AIDS
UPC	unique participant code
VDRL	venereal disease research laboratory
VL	viral load
WB	western blot
YCS	Y chromosomal sequences

Table of Contents

Foreword 1

Preface 2

1. Introduction 11

A. Survey preparation 17

1 Stakeholders 18
- A-1.1 Rationale for including stakeholders 18
- A-1.2 Stakeholder contributions 18
- A-1.3 Stakeholder engagement process 20
- A-1.4 References 20

2 Planning considerations before starting a biobehavioural survey 21
- A-2.1 Reflect on the HIV epidemic 21
- A-2.2 Consider epidemiological factors when choosing target population for BBS 21
- A-2.3 Conduct an initial review of information about proposed target population 22
- A-2.4 Ensure adherence to ethical standards 23
- A-2.5 Ensure funding is sufficient to meet objectives 23
- A-2.6 Consider timing and intervals between BBS rounds 23
- A-2.7 References 24

3 Ethical consideration in planning and conducting surveys 25
- A-3.1 International guidelines for human subjects research 25
- A-3.2 Fundamental ethical principles 26
- A-3.3 Research ethics committees 26
- A-3.4 Sampling of participants 26
- A-3.5 Obtaining voluntary informed consent 26
- A-3.6 Ensuring privacy and confidentiality 28
- A-3.7 Personal identifying information 29
- A-3.8 Data security 29
- A-3.9 Reimbursement to participants 29
- A-3.10 Return of test results 29
- A-3.11 Treatment and referral 29
- A-3.12 Adverse events, protocol deviations and protocol amendments 30
- A-3.13 Research ethics training 30
- A-3.14 Special considerations for incarcerated populations 30
- A-3.15 Web-based surveys 31
- A-3.16 References 31

4 Formative assessment 32
- A-4.1 The purpose of formative assessment 32
- A-4.2 Formative assessment objectives 33
- A-4.3 Conducting a formative assessment 35
- A-4.4 Optimal use of information gathered through the formative assessment 35

A-4.5 Ongoing formative assessment 35
A-4.6 References 35

5 Questionnaire development 36
A-5.1 Key steps in questionnaire development 36
A-5.2 Summary 43
A-5.3 References 43

6 Biomarker considerations 45
A-6.1 Biomarker selection 45
A-6.2 Resource availability 47
A-6.3 Ethical considerations 47
A-6.4 Providing test results to participants 47
A-6.5 On-site or off-site testing 48
A-6.6 Treatment and referral 48
A-6.7 Testing for HIV biomarkers 48
A-6.8 Testing for non-HIV biomarkers 50
A-6.9 Selection of tests and testing algorithms 54
A-6.10 Development of standard operating procedures 55
A-6.11 Specimen type, collection, labelling, processing and transport 55
A-6.12 Laboratory data management 59
A-6.13 Laboratory procedures 59
A-6.14 Specimen repository 59
A-6.15 Quality assurance (assessment) and quality control 59
A-6.16 Supply chain considerations 60
A-6.17 Biosafety and biosecurity 60
A-6.18 References 60

7 Preparing biological data-collection instruments 62
A-7.1 Specimen tracking forms 63
A-7.2 Specimen transport logs 63
A-7.3 Refrigerator or freezer temperature logs 63
A-7.4 Test results form 63

8 Eligibility for participation 65
A-8.1 Defining the population 65
A-8.2 Verifying eligibility 66
A-8.3 Elements of eligibility 66
A-8.4 Exclusion criteria 67

9 Sampling strategy 68
A-9.1 Concepts used in sampling, and types of sampling 68
A-9.2 Nonprobability sampling methods 72
A-9.3 Cluster sampling overview 74
A-9.4 Procedures for two-stage and multistage sampling 76
A-9.5 Conventional cluster sampling versus time-location cluster sampling 79
A-9.6 Time-location sampling 83
A-9.7 Respondent-driven sampling 92
A-9.8 References 101

10 Sample size requirements **103**
A-10.1 Considerations for determining sample size 103
A-10.2 Additional considerations 107
A-10.3 References 108

11 Population size estimation methods used with surveys **109**
A-11.1 Population size estimation based on surveys 112
A-11.2 Emerging methods 116
A-11.3 Selecting a method 117
A-11.4 References 119

12 Supplemental studies **121**
A-12.1 Qualitative studies 121
A-12.2 Cohort studies 130
A-12.3 Partner surveys 131
A-12.4 References 133

13 Data management **135**
A-13.1 Data documentation 135
A-13.2 Data dictionary 135
A-13.3 Unique participant identifier 136
A-13.4 Data-quality checks 136
A-13.5 Skip patterns 137
A-13.6 Data entry 137
A-13.7 Data confidentiality 137
A-13.8 Data access and use 137
A-13.9 Data backup 138
A-13.10 Data archiving and version control 138
A-13.11 Data security and storage 138

14 Staff selection **140**
A-14.1 Considerations for field staff selection 140
A-14.2 Field staff roles and responsibilities: non-method specific 141
A-14.3 Field staff roles and responsibilities: method specific 142
A-14.4 Considerations for staff payment 142

15 Standard operating procedures **143**
A-15.1 How to write a standard operating procedure 143
A-15.2 BBS activities that require an SOP 143
A-15.3 Sampling-specific SOPs 144
A-15.4 SOPs for checklists, logbooks labels and forms 144

16 Training and training methods **146**
A-16.1 Training 146
A-16.2 Training methods 147

B. Survey implementation and quality assurance 148

1 Data collection **149**
B-1.1 Surveys using nonprobability sampling methods 149
B-1.2 Conventional cluster sampling 152
B-1.3 Implementation of time-location sampling 156
B-1.4 Respondent-driven sampling 167

2 Survey monitoring **173**
B-2.1 Observing data collection 174
B-2.2 Laboratory monitoring 174

3 Preparing data for analysis **176**
B-3.1 Identifying and correcting data errors 176

C. Data analysis and use 185

1 Data analysis **186**
C-1.1 Developing a data analysis plan 186
C-1.2 Preparing a dataset for analysis 187
C-1.3 Analytical methods for BBS data 193
C-1.4 Combining multiple surveys and subsetting data for analysis 196
C-1.5 Comparing surveys 196
C-1.6 Suggested resources 199
C-1.7 References 199

2 Use and dissemination of survey findings **200**
C-2.1 Potential uses for BBS data 200
C-2.2 Identifying a target audience 201
C-2.3 Tailoring a message 202
C-2.4 Determining a dissemination strategy 202
C-2.5 Additional considerations for BBS data interpretation 203
C-2.6 Relevant websites 204
C-2.7 References 204

3 Evaluating survey implementation **205**
C-3.1 Data collection for evaluation 205
C-3.2 Using evaluation findings 206

4 Data sharing for public use **207**
C-4.1 Data Sharing 207
C-4.2 References 208

Supplemental Materials

Section I: Appendices 1

I-1: Protocol Checklist 2
I-2: Sample Budget Calculator and Gantt Chart (Survey Timeline) 10
I-3: Sample Consent Form 30
I-4: Survey Design Checklist 33
I-5: Brief Demographic Survey for Key Informants and Focus Groups 39
I-6: Formative Interview Guide for Gatekeepers 41
I-7: Formative Assessment Questionnaire Guide for Target Population 46
I-8: Formative Mapping 59
I-9: Safety and Security Guide 66
I-10: Specimens Used for HIV Testing 63
I-11: Material Transfer Agreement 69
I-12: Interpretation of Hepatitis Serologic Test Results 72
I-13: Specimen Tracking Form 74
I-14: Specimen Transportation Driver Log 76
I-15: Comparison of HIV Testing Technologies 78
I-16: Specimen Temperature Monitoring Logs 80
I-17: Sample HIV Rapid Testing Results Log 84
I-18: Equal Probability Sampling (EPS) and Probability Proportional to Size (PPS) Cluster Sampling Guide 86
I-19: Survey Implementation & Quality Assurance – Data Collection - TLS 91
I-20: Venue Observation Forms 94
I-21: Unique Participant Code (UPC) 99
I-22: Sample Size Requirements to Detect Differences 100
I-23: Calculator for the Number of Unique Objects for Estimating Population Size 101
I-24: Qualitative Interview Guide 102
I-25: Qualitative Interview for Target Population (FSW) Consent Form 109
I-26: Formative Assessment Consent Form for Key Informants 110
I-27: Data Management SOP for RDS Survey Using QDS 112
I-28: Sample Data Dictionary 114
I-29: Staff Data Agreement 116
I-30: Specimen Handling (Collection, Processing, and Storage) SOP 118
I-31: Recruitment Training Script 120
I-32: RDS Participant Checklist 122
I-33: Data Error Audit Log 125
I-34: Sample Table of Contents for Survey Report 126
I-35: Sample Press Release 128
I-36: Data Presentation Guide 130
I-37: Data Use (Confidentiality) Agreement 132

Section II: Questionnaire Modules 133

1. Questionnaire Parameters 136
2. Eligibility 140
3. RDS Recruiter-Recruit Relationship 151
4. RDS Personal Network Size 154
5. Size Estimation 156
6. ACASI Tutorial 159
7. Demographics 162
8. Transgender Characteristics 171

9. Sex Work Characteristics 176
10. Clients Of Sex Workers Characteristics 185
11. Incarceration Characteristics 188
12. Transport Worker Characteristics 196
13. Reproductive Health 200
14. Sexual History 208
15. Recall Sexual Behavior 218
16. Partner Concurrency 239
17. Last Sex Act 243
18. Condom Use/Accessibility 249
19. Lubricant Use 256
20. Microbicides 261
21. Alcohol Use 264
22. Non-Injection Drug Use 270
23. Injection Drug Use 277
24. Drug Overdose 295
25. Sexually Transmitted Infections 301
26. Hepatitis B and C 306
27. Services Uptake 309
28. PEP and PrEP 341
29. Shame, Stigma, Harassment & Discrimination 346
30. Physical Violence 369
31. Sexual Violence: Coercion and Rape 383
32. Internalized Homophobia 403
33. Depression 406
34. Suicide Ideation and Attempts 410
35. Social Support 415
36. Social Cohesion, Social Participation and Collective Agency 431
37. Game of Contacts 437
38. HIV Knowledge and Perceptions 439
39. Questionnaire Feedback/Interview Status 442
40. RDS, Peer Recruitment 444

Section III: Indicators 451

Introduction 454
What are indicators? 454
How are data collected for indicators? 455
How are indicators used? 456
Indicator selection for these guidelines 457

Indicators 460
Sexual debut 460
Condom use at last sex 461
Consistent condom use 463
Buying sex 464
Selling sex 465
Unprotected receptive anal intercourse 466
Alcohol use 467
Injecting drug use 470

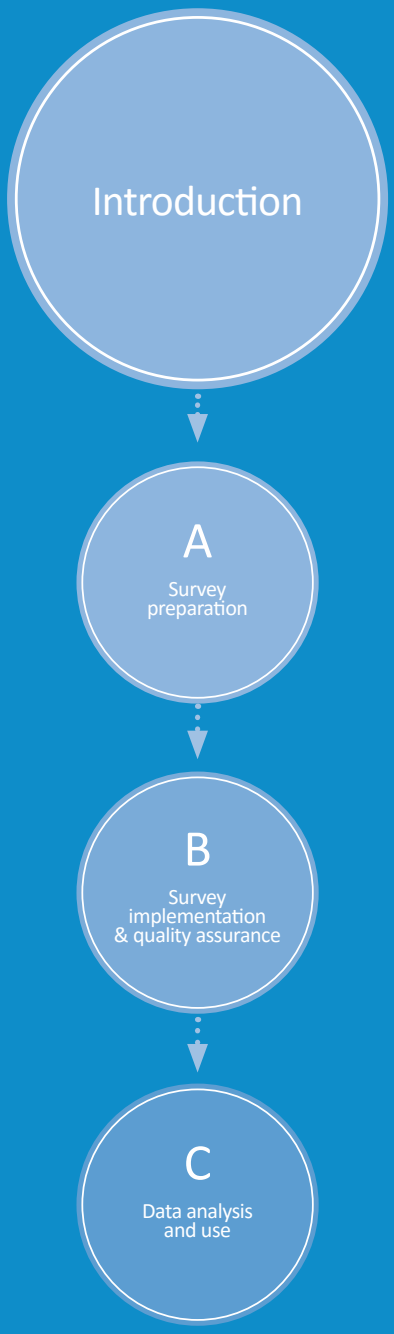
Received clean needles or syringes 471
Sharing injecting equipment 472
Sharing drug-preparation equipment 473
Use of sterile injecting equipment 474
Received medication for drug dependency 475
Drug overdose 476
Drug overdose prevention training 477
Targeted information, education and communications 478
Received condoms 479
Never tested for HIV 480
HIV test in last 12 months 481
Tuberculosis screening 482
Received tuberculosis treatment 483
Currently in care 484
CD4 count 485
Currently on antiretroviral therapy 486
Prevention of mother-to-child transmission 487
Hepatitis B virus testing 488
Hepatitis C virus testing 489
Discussed pre-exposure prophylaxis 490
Taken pre-exposure prophylaxis 491
Current use of modern family planning methods 492
Antenatal clinic access and HIV testing 493
Seeking care for sexually transmitted infection symptoms 494
Sexually transmitted infection screening 495
Received sexually transmitted infection treatment 496
Received lubricants 497
HIV knowledge and perceptions 498
Incarceration 499
Health-care stigma 500
Discrimination and social exclusion 501
Violence 502
Condom use at last sex (sex workers only) 504

Biomarker Indicators **505**
Eligible but not on HIV treatment 506
HIV treatment coverage 507
Viral load suppression 508
Unsuppressed viral load 509
Population viral load 510
HIV/sexually transmitted infection coinfection 511
HIV/hepatitis coinfection 512
HIV prevalence 513
Sexually transmitted infection prevalence 514

Section IV: Glossary **515**

Introduction

Background 1
Rationale for the guideline 2
Objectives 3
Target audience 4
Relationship to other resources 5
Structure 6
References 7



Background

An understanding of HIV burden, risk factors, and coverage of prevention and treatment services is critical for combatting the HIV epidemic *(1)*. Thus, biobehavioural surveys (BBS) assessing these parameters are integral components of a national HIV strategy and surveillance system. These guidelines outline the latest approaches and methodologies for planning and conducting such surveys. Countries that implement repeated BBS can monitor changes in their populations' risks for HIV, determinants of those risks, and access to prevention and treatment over time.

Many individuals at high risk for HIV are socially marginalized and may not identify themselves as such when accessing services. This makes it difficult to track them in HIV programme registers, and impedes efforts to assess the effectiveness of services. Such key populations include sex workers (SW), men who have sex with men (MSM), transgender women (TG) and people who inject drugs (PWID), all of whom are at increased risk for HIV infection compared with the population at large (*2-7*).

These guidelines describe sampling designs and methods for conducting surveys among key populations for which ready-made sampling frames are not typically available. As affirmed by the Joint United Nations Programme on HIV/AIDS (UNAIDS)/WHO in the publication *Guidelines for second generation HIV surveillance: an update: know your epidemic (8)* HIV surveillance among these key populations is a priority in all epidemic settings. In countries where the majority of HIV transmission takes place among members of the general population, data from HIV surveys among the general population should be used together with data from BBS among key populations to inform the epidemic response. This is due to the realization that countries previously categorized as having 'generalized' epidemics in fact experience 'mixed' epidemics, where HIV epidemics exist simultaneously among the general population and key populations. Separate surveys for key populations are warranted in such settings because key population members are usually not well identified or covered in general population surveys.

Although these guidelines focus on key populations, they are also applicable to other populations, such as clients of SW, prisoners, transportation workers and migrants. The term "target population" is used here to denote any survey group deemed to be at high risk for HIV infection, including key populations.

HIV control efforts directed at key and other high-risk populations can have a substantial impact on the epidemic. HIV policies and programmes will be more effective if they are informed by accurate measures of HIV prevalence and incidence among those populations, trends in their HIV-related risk behaviours, and the extent to which they access prevention and treatment services. This makes surveys critical tools for monitoring the HIV epidemic and evaluating HIV control efforts. UNAIDS and other groups have projected that 28–65% of new HIV infections in the People's Republic of China, the Dominican Republic, Kenya, Mozambique, Nigeria and Peru are among key populations and their partners *(2-7)*. Up to 30% of all HIV infections worldwide are due directly or indirectly to sex work, which can involve male, female and transgender individuals, demonstrating the importance of considering all genders when planning BBS *(9)*. Additionally, global prevalence of HIV among MSM is 19 times higher than in the general population, and prevalence among TG is 48 times higher *(10-12)*. Meanwhile, PWID represent an estimated 10% of all HIV infections globally *(13)* and are 22 times more likely to be HIV-infected compared to the general population *(14)*. For SW, HIV transmission is more likely to occur due to inconsistent condom use with multiple partners. MSM and TG are at high risk because HIV transmission is five times more likely to occur during unprotected, receptive anal sex than during vaginal sex *(15)*. Among PWID, sharing nonsterile needles or syringes is a highly efficient way to transmit HIV *(16)*.

Elevated risk for HIV transmission and acquisition among key populations is primarily related to four factors:

- efficient transmission of infection via unprotected anal sex and needle-sharing behaviours;
- frequent exposure to infection via multiple sex partners if not consistently using condoms;
- high HIV prevalence among networks of sexual or drug-injecting partners; and
- inferior access to quality health care compared to individuals in the general population.

Key Populations and Risk

Being an SW, MSM, TG or PWID alone does not increase the individual risk for getting or transmitting HIV. Rather, risk is elevated when individuals practise "unsafe" behaviours (i.e. have vaginal or receptive anal sex without a condom or share nonsterile injecting equipment) with partners who have a different, or discordant, HIV status from them.

The underlying reason that key populations are at higher risk for HIV is that HIV prevalence among their sexual or needle-sharing network is already high. Although members of the general population also engage in frequent vaginal and anal sex without condoms, their chance of having an HIV-infected partner is much lower than that of members of key populations.

Policies that criminalize behaviours of key populations or limit their access to services may further elevate their risk of acquiring and transmitting HIV. For example, more than two thirds of sub-Saharan African countries criminalize same-sex practices *(17)*. Criminalized populations receive less than 25% of intended human-rights-focused interventions, and may consequently lack access to services necessary for protecting themselves *(18)*. In countries where the behaviours of SW, MSM, TG or PWID are criminalized, and those laws are enforced, the negative impact on access to services is even greater *(19)*. Conversely, laws that prohibit discrimination against certain populations are correlated with better access to prevention services by the populations.

Sociostructural factors such as stigma, harassment, marginalization and lack of social support can contribute to higher transmission of HIV and other infections. This is particularly apparent among sex workers, sexual minorities and persons who inject drugs *(20, 21)*. Social marginalization or rejection can force individuals into isolation, unemployment, poverty and unstable living conditions or homelessness. It can also expose them to physical and sexual violence, and limit their access to HIV outreach commodities and educational messages *(22)*. Fear of police harassment, arrest or discrimination by health-care providers may also discourage members of key populations from accessing prevention and treatment services. In an assessment of the HIV prevention services available to MSM in 147 low- and middle-income countries, 55% were unable to document provision of even the most basic HIV-related services *(23)*. It is estimated that, worldwide, only 4% of PWID with HIV who are eligible for treatment are on antiretroviral therapy (ART) *(24)*.

Rationale for the guideline

In 2000, Family Health International (now known as FHI 360) issued a resource for HIV survey protocol development, implementation and data interpretation entitled *Guidelines for repeated behavioral surveys in populations at risk for HIV (25)*. Informally known as the "Red Book" because of the colour of its cover, that document has been widely used as a guide for conducting BBS among populations at risk for HIV.

The HIV landscape has changed significantly since the arrival of combination antiretroviral treatment in the 2000s in low- and middle-income countries. In addition, data needs and survey methods have changed substantially in the years since then. New sampling methods have emerged, such as respondent-driven sampling (RDS). Surveys increasingly include tests for biomarkers (e.g. HIV serostatus, CD4+ T-cell count, viral load, other sexually transmitted infections or viral hepatitis), qualitative interviews and population size estimation activities. The ethics of conducting surveys with HIV testing – particularly the return of test results to participants – have also evolved.

National governments, donors and United Nations (UN) agencies increasingly require data on access to HIV services, particularly because of the important role of antiretroviral therapy (ART) in HIV treatment and prevention. Engaging target populations and other stakeholders has become more important for survey planning and implementation, as has providing third-party researchers broader access to survey data.

These guidelines aim to facilitate BBS design and implementation in this context of evolving survey methods and measures. The ultimate purpose of the document is to facilitate the collection of high-quality survey data that will lead to informed public health action.

Objectives

These guidelines aim to:

- update the overall approach and methodology of BBS in light of advances made during the past decade;
- improve the quality of BBS by providing comprehensive guidance, particularly for sampling methods and the collection of interview and biomarker data;
- increase the relevance of BBS for public health programming by ensuring the collection of representative and actionable data; and
- promote the use of survey findings to improve service delivery, monitoring and evaluation, and policy development.

Target audience

These guidelines are written for policy-makers, organizations and technical staff who are planning or conducting BBS among key populations and other at-risk populations. The document is applicable for all settings, but particularly for low- and middle-income countries. We assume that the reader possesses a basic understanding of epidemiological and survey principles, sampling, inference and measures of risk.

Relationship to other resources

There are many excellent resources related to HIV surveys *(26)*, population size estimation *(27)*, and second-generation HIV surveillance *(1, 8)*. This document supplements and complements these resources.

However, BBS are just one data source informing key population HIV epidemics; others include population size estimation, mapping, HIV case surveillance, cohort studies, programme-based data and, further upstream, data on policy, stigma and discrimination. These guidelines are not intended for general population or household-based surveys; other resources are available for those purposes.

Structure

These guidelines are intended to provide an all-in-one resource for the conduct of BBS. Written in a concise manner, they focus on the "how to" and practical aspects of survey implementation – from formative assessment, survey design and protocol development to survey conduct, data analysis and dissemination of findings.

The core of the guidelines is divided into three principal sections, which are subdivided into chapters. Section A, ***Survey preparation***, is the largest part of the guidelines, and begins by discussing stakeholder involvement, formative assessment, human subjects considerations, survey instrument development, biomarker considerations, survey eligibility criteria, and population size estimation methods. Other chapters cover principal sampling strategies, sample size calculation, data management and data-collection procedures. Section A also covers practical aspects of survey preparation – including the development of standard operating procedures (SOP), ensuring adequate human resources and conducting appropriate staff training. Section B, ***Survey implementation and quality assurance***, focuses on monitoring and field supervision during survey implementation and data management, and specific cluster survey and RDS-related considerations for survey conduct. Section C, ***Data analysis and use***, addresses data analysis, data use, preparation of anonymized individual data files for data sharing, and evaluation of the completed survey.

The document also includes extensive supplementary materials: ***Appendices*** (Section I), ***Questionnaire modules*** (Section II), ***Indicators*** (Section III) and ***Glossary*** (Section IV). The collection of data through standardized data instruments is the core of any survey activity. The data instruments have undergone considerable modifications compared with those in the original guidelines; for example, they now include questions on the continuum of care. Rather than complete, stand-alone data instruments for distinct target populations, the instruments are organized as modules focused on particular topics. This provides users with more flexibility to construct data instruments according to local priorities and the needs of each target population. The data instruments' core variables have also been made available in ready-to-use electronic form.

Indicators and other standardized data measurements facilitate comparison of the HIV epidemic between different time points, locations or populations. The recommended reference sheets for data measurements are harmonized with indicators used by the UN and other global organizations engaged in HIV funding and programming. This section also includes additional data measurements recommended by the guidelines' authors.

The supplementary materials include an array of resources, including references for further reading and examples of protocol checklists, survey consent language and data analysis commands for commonly used software packages.

We hope these updated guidelines will facilitate the collection of more detailed, accurate and relevant data among populations at risk for HIV and thus facilitate more effective HIV services, advocacy and policy-making.

References

1 UNAIDS/WHO Working Group on Global HIV/AIDS and STI surveillance. Guidelines for second generation HIV surveillance: the next decade. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/WHO Working Group on Global HIV/AIDS and STI surveillance; 2000 (http://www.who.int/hiv/pub/surveillance/pub3/en/index.html, accessed 4 August 2016).

2 PAHO. Modos de transmisión del VIH en América Latina. Pan American Health Organization (PAHO); 2009 (http://www.unaids.org/en/media/unaids/contentassets/documents/countryreport/2009/20090810_MOT_Peru_es.pdf, accessed 3 August 2016).

3 UNAIDS. Country snapshot: China: HIV and men who have sex with men. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2012 (http://www.unaids.org.cn/pics/20121227100804.pdf, accessed 4 August 2016).

4 UNAIDS/COPRESIDA/DIGECITSS. HIV modes of transmission model: analysis of the distribution of new HIV infections in the Dominican Republic and recommendations for prevention. Santo Domingo, República Dominicana: Joint United Nations Programme on HIV/AIDS (UNAIDS), Consejo Presidencial del SIDA (COPRESIDA), Dirección General de Infecciones de Transmisión Sexual y SIDA (DIGECITSS); 2010 (http://www.unaids.org/en/media/unaids/contentassets/documents/countryreport/2010/201011_MOT_DominicanRepublic_en.pdf, accessed 3 August 2016).

5 UNAIDS/World Bank. New HIV infections by modes of transmission in West Africa: a multi-country analysis. Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Bank; 2010 (http://www.unaids.org/sites/default/files/en/media/unaids/contentassets/documents/countryreport/2010/201003_MOT_West_Africa_en.pdf, accessed 4 August 2016).

6 World Bank/UNAIDS. Kenya: HIV prevention response and modes of transmission analysis. World Bank/Joint United Nations Programme on HIV/AIDS (UNAIDS); 2009 (http://siteresources.worldbank.org/INTHIVAIDS/Resources/375798-1103037153392/KenyaMOT22March09Final.pdf, accessed 4 August 2016).

7 World Bank/UNAIDS. Analysis of modes of HIV transmission and national response to HIV and AIDS. Maputo, Mozambique: World Bank/Joint United Nations Programme on HIV/AIDS (UNAIDS); 2010 (http://regist2.virology-education.com/4thINTEREST/docs/14_Fazenda.pdf, accessed 4 August 2016).

8 UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. Guidelines for second generation HIV surveillance: an update: know your epidemic. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2013 (http://apps.who.int/iris/bitstream/10665/85511/1/9789241505826_eng.pdf, accessed 3 August 2016).

9 Baral S, Beyrer C, Muessig K, Poteat T, Wirtz AL, Decker MR et al. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Infect Dis. 2012;12(7):538–549.

10 Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000–2006: a systematic review. PLoS Med. 2007;4(12):e339 (http://www.ncbi.nlm.nih.gov/pubmed/18052602, accessed 4 August 2016).

11 Baral SD, Poteat T, Stromdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. Lancet Infect Dis. 2013;13(3):214–222.

12 Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL et al. Global epidemiology of HIV infection in men who have sex with men. Lancet. 2012;380(9839):367–377.

13 Mathers BM, Degenhardt L, Phillips B, Wiessing L, Hickman M, Strathdee SA et al. Global epidemiology of injecting drug use and HIV among people who inject drugs: a systematic review. Lancet. 2008;372(9651):1733–1745.

14 UNAIDS. Global report: UNAIDS report on the global AIDS epidemic 2012. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2012 (http://www.unaids.org/sites/default/files/media_asset/20121120_UNAIDS_Global_Report_2012_with_annexes_en_1.pdf, accessed 4 August 2016).

15 Varghese B, Maher JE, Peterman TA, Branson BM, Steketee RW. Reducing the risk of sexual HIV transmission: quantifying the per-act risk for HIV on the basis of choice of partner, sex act, and condom use. Sex Transm Dis. 2002;29(1):38–43.

16 Kaplan EH, Heimer R. A model-based estimate of HIV infectivity via needle sharing. J Acquir Immune Defic Syndr. 1992;5(11):1116–1118.

17 Itaborahy LP, Zhu J. State-sponsored homophobia: a world survey of laws: criminalisation, protection and recognition of same-sex love. International Lesbian Gay Bisexual Trans and Intersex Association (ILGA); 2013 (http://old.ilga.org/Statehomophobia/ILGA_State_Sponsored_Homophobia_2013.pdf, accessed 4 August 2016).

18 UNDP/The Global Fund/UNAIDS. Analysis of key human rights programmes in Global Fund-supported HIV programmes. New York: United Nations Development Programme (UNDP)/The Global Fund/Joint United Nations Programme on HIV/AIDS (UNAIDS); 2011 (http://www.undp.org/content/dam/aplaws/publication/en/publications/hiv-aids/analysis-of-key-human-rights-programmes-in-global-fund-supported-hiv-programmes/Analysis%20of%20Key%20HRTS%20Programmes%20in%20GF-Supported%20HIV%20Programmes.pdf, accessed 4 August 2016).

19 PAHO. Improving access of key populations to comprehensive HIV health services: towards a Caribbean consensus. Washington D.C.: Pan American Health Organization (PAHO); 2011 (http://www.paho.org/hq/index.php?option=com_docman&task=doc_download&gid=16974&Itemid=, accessed 4 August 2016).

20 Altice FL, Kamarulzaman A, Soriano VV, Schechter M, Friedland GH. Treatment of medical, psychiatric, and substance-use comorbidities in people infected with HIV who use drugs. Lancet. 2010;376(9738):367–387.

21 Rhodes T, Singer M, Bourgois P, Friedman SR, Strathdee SA. The social structural production of HIV risk among injecting drug users. Soc Sci Med. 2005;61(5):1026–1044.

22 Centers for Disease Control and Prevention. Integrated prevention services for HIV infection, viral hepatitis, sexually transmitted diseases, and tuberculosis for persons who use drugs illicitly: summary guidance from CDC and the U.S. Department of Health and Human Services. MMWR. 2012;61(RR05):1–40 (http://www.cdc.gov/mmwr/pdf/rr/rr6105.pdf, accessed 8 August 2016).

23 UNAIDS. The Prevention Gap Report. Geneva: UNAIDS; 2016. (http://www.unaids.org/sites/default/files/media_asset/2016-prevention-gap-report_en.pdf, accessed 31 August 2017)

24 Harm Reduction International (2016) 'Global State of Harm Reduction

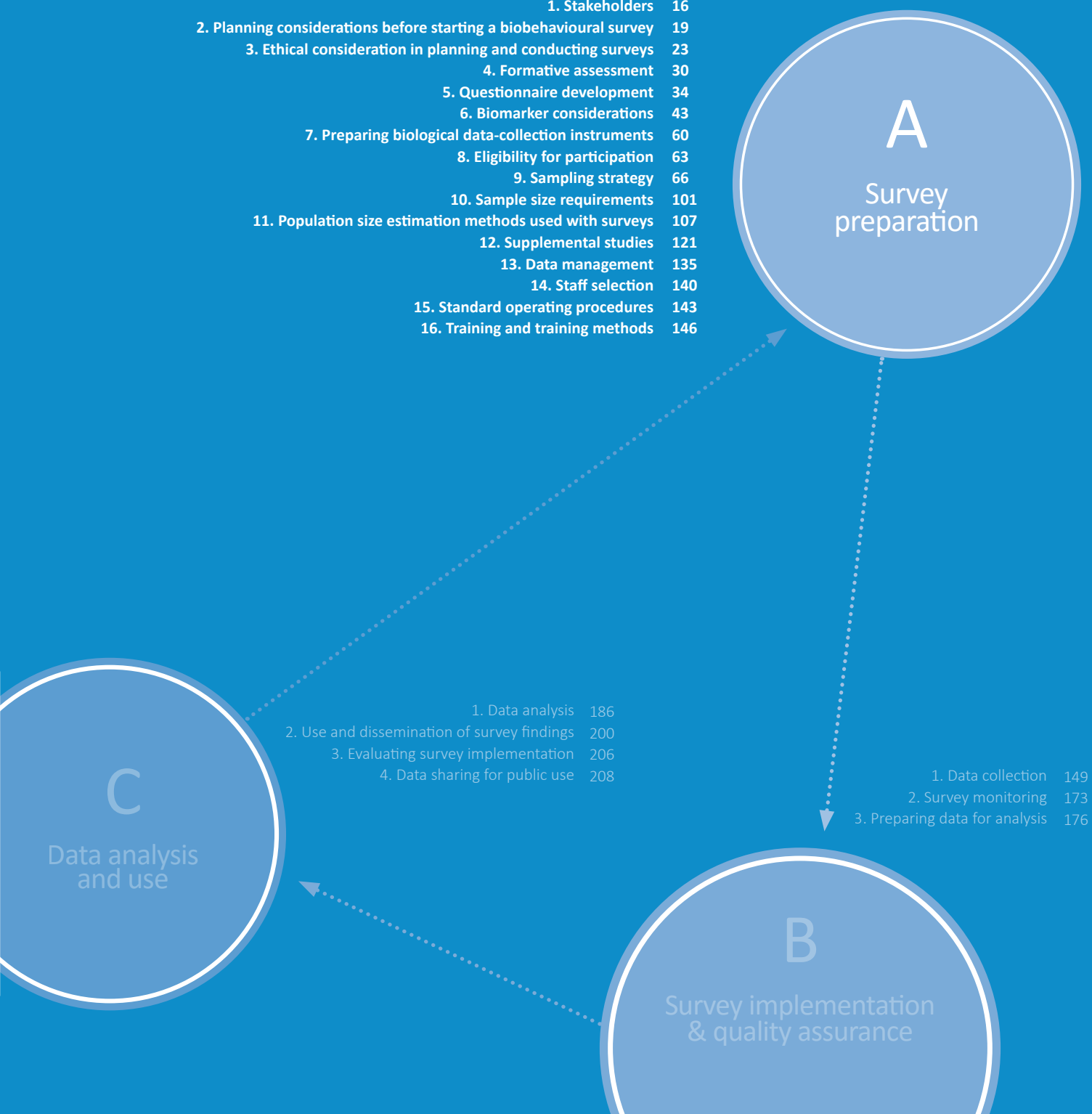
25 Amon J, Brown T, Hogle J, MacNeil J, Magnani R, Mills S et al. Behavioral surveillance surveys BSS: guidelines for repeated behavioral surveys in populations at risk of HIV. Arlington: Family Health International (FHI); 2000 (http://www.who.int/hiv/strategic/en/bss_fhi2000.pdf, accessed 3 August 2016).

26 UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. Guidelines on surveillance among populations most at risk for HIV. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2011 (http://www.unaids.org/en/media/unaids/contentassets/restore/20110518_Surveillance_among_most_at_risk.pdf, accessed 26 June 2016).

27 UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. Guidelines on estimating the size of populations most at risk to HIV. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2010 (http://www.unaids.org/sites/default/files/media_asset/2011_Estimating_Populations_en_0.pdf, accessed 4 August 2016).

A. Survey preparation

1. Stakeholders 16
2. Planning considerations before starting a biobehavioural survey 19
3. Ethical consideration in planning and conducting surveys 23
4. Formative assessment 30
5. Questionnaire development 34
6. Biomarker considerations 43
7. Preparing biological data-collection instruments 60
8. Eligibility for participation 63
9. Sampling strategy 66
10. Sample size requirements 101
11. Population size estimation methods used with surveys 107
12. Supplemental studies 121
13. Data management 135
14. Staff selection 140
15. Standard operating procedures 143
16. Training and training methods 146



Section A

1. Stakeholders

This chapter describes the role that stakeholders play in survey planning and preparation. Stakeholders are individuals or organizations with an interest in or concern for the implementation and results of a biobehavioural survey (BBS).

Stakeholders may include:

- agencies of the local or national government; for example, different ministries, provincial health authorities, national reference laboratories or HIV control programmes (as appropriate);
- development agencies;
- academic institutions;
- service providers;
- nongovernmental organizations (NGOs) and community-based organizations;
- donors;
- community leaders; and
- members of the target population.

Key terms

Community-based organization:	An organization largely composed of members from a given population (e.g. men who have sex with men), often working in advocacy or providing services.
Nongovernmental organization (NGO):	An organization founded by citizens and not part of the government and not a for-profit company.
Stakeholder:	A party or group of people with an interest in a given activity, who can contribute to it or is affected by it.

A-1.1 Rationale for including stakeholders

Various stakeholders can contribute to the successful implementation of BBS at every stage of the process. They can help to ensure that BBS design is consistent with on-the-ground realities, encourage target population participation, and provide resources or services to facilitate survey implementation. The primary consumers of the survey findings, including the national HIV programme and donors, should be involved in the planning and implementation of the BBS. Early stakeholder engagement often leads to greater use of BBS results to inform services, policy and advocacy.

Target population support for the survey is crucial, and investigators planning a BBS should engage influential members early in the process, both to seek their input on survey content and to communicate how the data may be used. Investigators can also learn about the local context from service providers including clinicians, NGOs and other civil society organizations that offer outreach services. Because of their close relationship with the target population, these stakeholders may also be able to encourage participation. Local academic institutions may be able to offer guidance on sampling strategies, data analysis and report development.

Investigators may also benefit from engaging legislators and politicians, who can allocate resources for surveys, and use the findings to advocate for policies that benefit the target population. Such stakeholders can also improve access to services and prevent harassment of the target population during the BBS.

Although stakeholders make important contributions, national officials should play a lead role in choosing the target populations, variables of interest and types of analyses, and in disseminating the data. Often, leadership by a national body increases the likelihood that the BBS will be a high-quality survey, will serve the needs of the country as a whole and will result in data that are widely used.

A-1.2 Stakeholder contributions

Individuals and groups may be involved in all aspects of survey planning and implementation, from financial support to dissemination of findings (see Table A-1.1). Stakeholder dialogue is essential when there may be a difference of opinion about the scope (e.g. survey location, data and biomarker measures) or objectives of a BBS. For example, stakeholders may have different views on the prioritization of one target group or the inclusion of biomarkers related or unrelated to HIV.

Dialogue among stakeholders can clarify these issues, and can be useful in considering options and making decisions. Common stakeholder expectations are essential for a successful survey. Stakeholder contributions might include:

- defining the scope of the survey;
- defining the populations to be surveyed;
- providing input on the data and biomarkers to be collected;
- raising funds for the survey;
- promoting participation in and acceptance of the survey; and
- employing survey findings to develop and advocate for policies.

Table A-1.1 Potential stakeholders to consider for BBS planning and implementation activities

Stakeholder	What they offer
• Members of the target population • Community leaders	• Ensure the survey is perceived as legitimate by the target population • Provide investigators with social context • Provide information on the needs of target populations • Provide insight into: - survey content - appropriateness of sampling strategy - survey design - wording of the questionnaire
• Nongovernmental organizations • Community-based organizations	• Provide social context • Encourage participation by the target population • Provide information on the needs of target populations • Facilitate collaboration with the target population
• Ministry of Health • National HIV/AIDS programme or committee	• Ensure investigators have the necessary approvals from the national, provincial and local government • Provide legal and social context • Provide input on survey design • Coordinate with existing government services for target populations • Ensure that different government entities (including law enforcement, clinics and laboratories) are informed of the survey, as appropriate, and do not obstruct survey implementation • Integrate surveillance activities into the national monitoring and evaluation plan • Minimize risk that efforts are duplicated by different organizations • Disseminate results
• HIV service providers and policymakers	• Provide information on the needs of target populations • Ensure that survey results provide actionable information on a timely basis • Guide strategic planning and advocacy • Facilitate biological specimen collection, cold chain and storage
• Donors	• Provide funding; formulate information needs
• Academic institutions	• Provide guidance on sampling strategies
• Research organizations	• Support survey implementation and data analysis • Draw on experience implementing similar surveys in the past
• Staff engaged in surveillance for sexually transmitted infections (STIs)	• Provide information on the needs of target populations • Facilitate collaboration with the target population • Improve efficiency by integrating STI and HIV surveillance activities and survey findings • Ensure appropriate biomarkers are collected
• Police • Ministry of Security • Correctional officers (jails/prisons)	• Depending on country or local context, ensure the safety and security of surveyors and participants, • For incarcerated populations, ensure proper ethical protections and procedures are in place

Source: UNAIDS/WHO 2011 *(1)*

A-1.3 Stakeholder engagement process

It may be helpful to hold regular consultations with stakeholders throughout all stages of the BBS, to ensure that they contribute ideas and exchange experiences. This can be achieved by forming a working group of stakeholders that meets regularly. The group can convene to review key areas such as the scope of the survey, methods for implementation, and the plan for disseminating and employing the resulting data *(2)*. These meetings may result in changes to questionnaires, testing approaches or methods that are best explored during the formative assessment or before the BBS is implemented. In some situations, it may be necessary to coordinate between stakeholder groups with different areas of interest, such as budget and technical stakeholder groups.

A-1.4 References

1 UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. Guidelines on surveillance among populations most at risk for HIV. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2011 (http://www.unaids.org/en/media/unaids/contentassets/restore/20110518_Surveillance_among_most_at_risk.pdf, accessed 26 June 2016).

2 Amon J, Brown T, Hogle J, MacNeil J, Magnani R, Mills S et al. Behavioral surveillance surveys BSS: guidelines for repeated behavioral surveys in populations at risk of HIV. Arlington: Family Health International (FHI); 2000 (http://www.who.int/hiv/strategic/en/bss_fhi2000.pdf, accessed 3 August 2016).

2. Planning considerations before starting a biobehavioural survey

Before committing to conducting a biobehavioural survey (BBS), investigators should establish what is already known about the HIV epidemic in the country and within the target population, and assess whether a BBS is warranted. This initial data review should serve to quickly gather and synthesize preliminary information justifying the BBS; it need not be extensive. This chapter describes the process for conducting this initial review.

Key terms

Risk:	In ethical terms, the exposure to injury, loss or harm, expressed in terms of the probability and magnitude of that harm. Risks to subjects must be minimized and must be reasonable in relation to anticipated benefits to subjects and the importance of the expected knowledge.
Time-location sampling (TLS):	Recruits individuals (target population members) from specific locations (venues) during specific time periods; thus, TLS surveys represent only those who frequent specific venues.
Vulnerability:	The (relative) powerlessness to resist or survive the effects of an adverse environment.

A-2.1 Reflect on the HIV epidemic

Before conducting a BBS of key populations – sex workers (SW), men who have sex with men (MSM), transgender women (TG) and people who inject drugs (PWID) – it is best to conduct a review, to ensure that the BBS will not duplicate efforts and to confirm that it is feasible. A lack of data on other populations does not imply that a BBS should be conducted; other factors should be taken into account, including risk behaviour, ability to reach the sample size and HIV prevalence.

The methods presented below are not intended as a replacement for discussions with stakeholders. Rather, they are a means for producing an evidence base to guide decisions about conducting a BBS. This evidence should be reviewed with all relevant stakeholders, including members of the proposed target population, before making a final decision about the BBS.

A-2.2 Consider epidemiological factors when choosing target population for BBS

When considering a BBS with a target population, investigators should first consider the role that the group plays in the local HIV epidemic. This may be done by examining risk factors, vulnerability, population size and HIV prevalence, as discussed below.

A-2.2.1 Risk factors

Most BBS are conducted among populations who are defined by their risk factors and behaviours (e.g. commercial sex, anal sex or injecting drug use), or because investigators suspect that the population exhibits risk behaviours more frequently than the general population. For target populations other than key populations, it is important to assess what is known about the behaviours of those populations before committing to a BBS of them. For example, where data on clients of SW are not available, truck drivers have been studied as a proxy, although not all truck drivers may be clients of SW. Although not representative of all clients, surveys of truck drivers may shed light on HIV transmission dynamics along truck routes. Before conducting a survey on truck drivers, it is important to gather available information on the drivers and how their HIV-related risk behaviours resemble or differ from those of other populations.

A-2.2.2 Vulnerability

Vulnerability refers to factors typically outside an individual's control that indirectly raise an individual's risk for HIV infection – including unequal opportunities, social exclusion, unemployment or precarious employment *(1)*. Other populations are deemed vulnerable due to

poor access to prevention, care and treatment services. A target population may be of interest because of its vulnerabilities, although these may only be apparent in the context of other groups. For example, migrant workers can experience social exclusion, arrest and unequal employment opportunities. In contrast, there is little reason to think that teachers face vulnerabilities that make them more susceptible to HIV; in some cases, teaching may even represent a position of power.

A-2.2.3 Population size

The target population needs to be sufficiently large to ensure an adequate sample size can be reached. As a general rule, a BBS intended to be representative of a target population must have a sample of at least 500 respondents. Thus, it is not possible to conduct a BBS in a small community. A town of 5000 people is unlikely to have 500 MSM; even if it did and the survey team could identify all of them, it is unlikely they would all be eligible for the survey or willing to participate.

A-2.2.4 HIV prevalence

HIV prevalence indicates the proportion of individuals in a population who have HIV at a given time *(1)*. Generally speaking, BBS should be prioritized for populations with the highest HIV prevalence and impact on the overall epidemic. HIV prevalence is generally higher among SW, MSM, TG and PWID in mixed and concentrated epidemics alike *(2-7)*. However, BBS on other populations may also be warranted if the data review suggests that they also have an elevated prevalence of HIV.

A-2.3 Conduct an initial review of information about proposed target population

Investigators can begin by reviewing published and unpublished information about the target population. Report summaries and journal abstracts can be accessed free of charge online.[1] For countries and populations where reports do not exist, literature from neighbouring countries or similar populations may provide useful insights.

Investigators can also review service data; such data are especially important if there are few or no reports or journal articles about the local target population. HIV testing data can provide valuable information, especially if the testing facility serves specific populations.

Information should also be sought from stakeholders including officials from the ministry of health, donors, health-care workers, laboratory officials, NGOs and community-based organizations that work with the population, other groups that interact regularly with them and, most importantly, members of the proposed target population. For example, brothel managers may know whether truck drivers access SW. Health-care workers may be able to describe their clientele and indicate if the proposed target population has a high burden of HIV.

Once sufficient information has been gathered, investigators can synthesize it and share the findings with other stakeholders. The group can weigh the factors described above and determine the value of conducting a BBS with the proposed target population.

A-2.3.1 Consider security concerns for the target population and the survey team

Where it is determined that a BBS with a given population will be valuable, investigators must then consider the security dimensions of survey implementation; for example, the survey may draw undue attention to the entire target population. Investigators should weigh the potential for risk to participants and the larger target population as a whole against the benefits of the survey,[2] and the safety of the survey team and participants should be a top priority in planning and implementation. If the threat of danger to the survey team and participants is agreed to be too high after consulting with local experts and community-based organizations, then it may be necessary to cancel the survey. If the survey is conducted, then safety procedures must be in place to reduce the chances of placing staff and participants in danger, to reduce the harm that could be done to staff and participants, and to respond quickly and appropriately in case a safety issue does occur. These procedures should include plans for dealing with hostile visitors or participants, reporting incidents through chain of command, conducting a pre-survey checklist for safety, emergency security actions and data security.

A-2.3.2 Develop a survey protocol

Next, the survey protocol and consent forms must be developed. A protocol specifies the survey objectives, defines the population that will be surveyed and the

[1] For example, see www.unaids.org and www.pubmed.gov.

[2] Information on human rights is available at: http://www.who.int/gender-equity-rights/en/, and information on health and human rights is available at http://www.who.int/hhr/HHR%20linkages.pdf?ua=1).

geographical area in which it will be conducted, and describes all the procedures involved (including sampling, obtaining consent, conducting interviews and providing counselling, collecting and testing biological specimen, and managing and analysing data) *(8)*. At this time, to ensure that the survey objectives are satisfied when the data are collected and later analysed, investigators should develop comprehensive plans for data management (Chapter A-13) and data analysis (Chapter C-1) along with the survey protocol. Guidance and a checklist for developing a protocol are included (Appendix I-1). When developing the protocol it is important to consider the following:

- what questions about the target populations need to be answered?
- what information is needed to address these questions?
- should the size of the target population be estimated?
- should a supplemental study be conducted (e.g. a qualitative, cohort or partner study)?
- what biomarkers should be collected?
- how will the data be used?

Suggested key objectives for biobehavioral surveys include:

- to estimate the proportion of suppressed HIV viral load (among people living with HIV)
- to estimate the prevalence of HIV infection
- to estimate access to and uptake of HIV prevention, care and treatment services
- to examine correlates of HIV infection
- to estimate the population size

A-2.4 Ensure adherence to ethical standards

Participants who are afraid to share information about risk behaviours or HIV/sexually transmitted infection (STI) status may drop out of the survey or falsify their responses. To successfully survey a marginalized population, these biases must be minimized by obtaining fully informed consent from participants and ensuring their absolute confidentiality *(9)*. Therefore, all survey protocols should be assessed by ethical review boards or human subjects oversight committees to ensure that they adhere to national and international laws and ethical standards *(10)*. More information on ethical considerations of survey conduct and the human subjects review process can be found in Chapter A-3.

A-2.5 Ensure funding is sufficient to meet objectives

The scope of a BBS is dependent on the financial resources available. Core budget components usually include:

- administrative costs, including overhead and management;
- staff costs, including investigators, data collectors, administrative support team and drivers;
- laboratory consumables and equipment costs, including gloves, specimen collection instruments and assays;
- travel costs, including transportation to the survey site if that is not in the same city as the investigators, or international technical advisors or investigators; and
- documentation and dissemination-related expenses, including report writing, editing, layout, and national and sub-national disseminations.

When estimating a BBS budget, important considerations include the target sample size, duration of survey, number of survey staff and biomarkers on which data will be collected. Investigators should estimate the number of HIV-infected participants and the follow-up tests that will be provided for each (e.g. CD4 and viral load). Costs will also vary by survey design; that is, conventional cluster sampling (CCS), respondent-driven sampling (RDS) and time-location sampling (TLS). A budget template to facilitate resource planning and alignment with the protocol can be found in Appendix I-2.

A-2.6 Consider timing and intervals between BBS rounds

Ideally, investigators should conduct BBS at intervals in order to capture changes in risk behaviours and the HIV/STI epidemic over time. The Joint United Nations Programme on HIV/AIDS (UNAIDS)/WHO *Guidelines for second generation HIV surveillance recommends* conducting a BBS every 1–3 years in key populations and every 3–5 years in other target populations, as warranted *(9)*. One year is not typically long enough for changes in HIV/STI prevalence to be detected. Depending on survey objectives, local resources and context, investigators are advised to consider waiting 18–24 months in order to yield more practical results.

Regardless of the time interval between BBS rounds, the benefits of regular data collection include:

- information for continuously reevaluating service coverage and needs;
- comparison estimates (or trend data, if three or more rounds of BBS have been implemented) for examining the effects of interventions over time; and
- the development of survey implementation and management capacity among national and local institutions.

Trend data from repeated BBS yield useful information for monitoring the epidemic. Although methodological consistency is important to produce these data, investigators should make advisable changes to variables, assays or methods where necessary rather than seeking to maintain the same instruments over time. It is possible to make such changes to improve the relevance or quality of survey results and still preserve trend analyses.

Circumstances may exist where BBS or subsequent rounds of surveying are not advised. For example, national priorities may shift away from key populations, limited financial or human resources may be devoted to other public health priorities, or the HIV/STI epidemic or patterns of risk behaviours may change. The decision to stop conducting BBS of a population must be considered in the context of the local epidemiological, political and financial environment. In circumstances where the HIV prevalence of the target population is lower than that of the general population, or where funding is no longer available, it may not be appropriate or feasible to continue BBS.

A-2.7 References

1 UNAIDS. UNAIDS terminology guidelines. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2011 (http://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2118_terminology-guidelines_en.pdf, accessed 4 August 2016).

2 Baral S, Beyrer C, Muessig K, Poteat T, Wirtz AL, Decker MR et al. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Infect Dis. 2012;12(7):538–549.

3 Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000–2006: a systematic review. PLoS Med. 2007;4(12):e339 (http://www.ncbi.nlm.nih.gov/pubmed/18052602, accessed 4 August 2016).

4 Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL et al. Global epidemiology of HIV infection in men who have sex with men. Lancet. 2012;380(9839):367–377.

5 Mathers BM, Degenhardt L, Phillips B, Wiessing L, Hickman M, Strathdee SA et al. Global epidemiology of injecting drug use and HIV among people who inject drugs: a systematic review. Lancet. 2008;372(9651):1733–1745.

6 UNAIDS. Global report: UNAIDS report on the global AIDS epidemic 2012. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2012 (http://www.unaids.org/sites/default/files/media_asset/20121120_UNAIDS_Global_Report_2012_with_annexes_en_1.pdf, accessed 4 August 2016).

7 WHO. Prevention and treatment of HIV and other sexually transmitted infections for sex workers in low-and middle-income countries: recommendations for a public health approach. Geneva, World Health Organization (WHO). 2012 (http://apps.who.int/iris/bitstream/10665/77745/1/9789241504744_eng.pdf, accessed 25 August 2016).

8 CDC. Developing a protocol. Atlanta, GA: Centers for Disease Control and Prevention (CDC); 2013 (http://www.cdc.gov/globalhealth/healthprotection/fetp/training_modules/13/developing-protocols_pw_final_09262013.pdf, accessed 1 November 2016).

9 UNAIDS/WHO Working Group on Global HIV/AIDS and STI surveillance. Guidelines for second generation HIV surveillance: the next decade. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/WHO Working Group on Global HIV/AIDS and STI surveillance; 2000 (http://www.who.int/hiv/pub/surveillance/pub3/en/index.html, accessed 4 August 2016).

10 WHO/UNAIDS. Guiding principles on ethical issues in HIV surveillance. Geneva: World Health Organization (WHO)/Joint United Nations Programme on HIV/AIDS (UNAIDS); 2013 (http://www.who.int/hiv/pub/surveillance/2013package/module2/en/, accessed 3 August 2016).

3. Ethical consideration in planning and conducting surveys

This chapter describes the ethical considerations of survey conduct and the human subjects review process. As with any research on people (also known as human subjects), biobehavioural surveys (BBS) have the potential to harm the individuals involved. BBS often involve members of marginalized, criminalized or otherwise vulnerable populations, and may collect sensitive data about stigmatized issues and behaviours.

Key terms

Anonymous: Describes something having no known name or identity. Achieved by removing all personal identifying information from an interview record or biological specimen.

Assent: An agreement by persons who cannot legally agree (consent) to survey participation (e.g. minors). Usually, consent by parents or caregivers is necessary in addition to assent by a minor potential survey participant.

Autonomous: Freedom from external control or influence; independence.

Beneficence: The effort to secure the well-being of survey participants, both by avoiding or minimizing harm and by maximizing possible benefits.

Confidential: The expectation or promise that information will be protected from disclosure to others.

Consent: Permission given by an individual to proceed with a specific test or procedure, with an understanding of the risks, benefits, limitations and potential implications of the procedure itself and its results.

Fingerprint codes: Fingerprint scanners, together with the appropriate software, convert a finger print image into an alphanumeric code without storing the actual image. If such codes cannot be used to reconstruct the fingerprint images they may be regarded as anonymous ID numbers. These codes are anonymous, but can nevertheless be used to link different visits of the same participants.

Justice: The fair selection of research participants and the fair distribution of risks and benefits for the study participants.

Such surveys must be designed and conducted ethically to ensure that participants can freely give informed consent, that their privacy and confidentiality are protected, and that the survey's benefits outweigh the potential harm to participants. The international community has developed ethical principles and guidance for the protection of human subjects that are applicable to both experimental research and surveys. Investigators and staff, who take responsibility for the well-being and safety of survey participants, must follow these ethical principles throughout the conducting of a survey.

A-3.1 International guidelines for human subjects research

International guidelines for human subjects research have been published over the past century in a handful of key documents. As the direct result of reported abuses of human subjects in biomedical experiments during the Second World War, the international community drafted the Nuremburg code (1947) *(1)* – a set of 10 requirements for the legitimate conduct of experimental research involving human subjects. The Declaration of Helsinki (published in 1964, last updated in 2013) *(2)* was primarily developed to guide physicians, but is considered by many to be the first global standard for ethical principles for human subjects research. *The Belmont Report, Ethical principles and guidelines for the protection of human subjects of research* (1979) *(3)*, published by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, lays out three fundamental principles for the ethical conduct of human subjects research: respect for persons, beneficence and justice. The International *Ethical guidelines for biomedical research involving human subjects* (published in 1982, last updated in 2002) *(4)* describes how ethical principles can be applied in practice, particularly in developing countries. UNAIDS/ WHO also published the *Guiding principles on ethical issues in HIV surveillance* in 2013 *(5)*.

A-3.2 Fundamental ethical principles

Respect for persons

"*Respect for persons*" is the principle that all individuals should be treated as autonomous (i.e. capable of making informed decisions) and their choices should be respected. For potential participants to make informed decisions to participate in a survey, survey staff must adequately inform them about the survey's objectives, its possible risks and benefits, and their right to refuse to be involved in or to withdraw from the survey at any time without penalty. Survey participants should decide whether to participate voluntarily, without undue influence or coercion or consequence, and with full knowledge of the procedures involved in participation.

Vulnerable persons may be unable to make informed decisions or be at a higher risk of being coerced. Studies involving such individuals must offer extra protections or precautions to ensure that their decision to participate is made voluntarily. Examples of vulnerable persons include individuals whose behaviours are stigmatized or criminalized – for example, people living with HIV, men who have sex with men (MSM), sex workers (SW), people who inject drugs (PWID) and transgender persons (TG) – and minors, those with mental or behavioural disabilities, subordinate members of hierarchal groups, elderly persons, pregnant women, incarcerated adults, ethnic minorities and displaced persons.

Beneficence

"*Beneficence*" is the principle that the goal of the research should be the welfare of the subject. The researcher should strive to "*do no harm*" and to maximize possible benefits while minimizing possible harms. Ethically, surveys can only be conducted if the public health benefit is likely to outweigh the personal risks of the participants involved. Investigators are responsible for participants' physical, mental and social well-being throughout their participation in the survey. They should be open with potential participants about the risks, discomforts and burdens that may be involved in the survey, and monitor and seek to minimize these factors throughout the process.

Justice

"*Justice*" is the principle that the risks and benefits of the research should be shared by the individual participant and the larger community. Investigators must ensure that the selection of survey participants is equitable (fair) and nonexploitive, meaning that participants are not selected because of their easy availability, compromised position (vulnerability) or manipulability (i.e. can be tricked into survey participation).

A-3.3 Research ethics committees

Research ethics committees (RECs), sometimes known as institutional review boards (IRBs), are responsible for reviewing research and surveys, and ensuring that they are designed and conducted in accordance with the three fundamental ethical principles. Investigators should always submit their survey protocol to an REC for review and approval before survey implementation begins.

During protocol review, RECs should ensure that:

- the survey design and conduct support the safety and well-being of participants;
- the recruitment strategy is appropriate and equitable;
- informed consent processes meet ethical standards and are sufficiently described;
- the possible benefits to the target population outweigh the risks;
- risks to participants are monitored and minimized; and
- privacy and confidentiality are protected.

A-3.4 Sampling of participants

Investigators should sample participants in a way that both protects participants and respects ethical principles. To help avoid unintentionally coercing individuals into joining the survey, investigators should assure potential participants that there is no penalty for declining participation. Investigators should also recruit participants in a way that separates or minimizes the link between survey participation and any services the potential participant is eligible to receive. Reimbursement for time should be reasonable (i.e. not too high). Where possible, survey-related benefits (e.g. free serological testing with results provided) should be offered to potential participants even if they decline to participate.

A-3.5 Obtaining voluntary informed consent

"Voluntary informed consent" occurs when an autonomous person with a good understanding of the survey actively decides to participate in the survey without being influenced or pressured. Consent is a process that starts before survey participation and continues until participation is complete. It is the investigator's responsibility to ensure that the potential participant receives enough information to give voluntary informed consent. Information should be provided in a manner that is appropriate for the culture, language ability, education level and general contextual environment (see Appendix I-3).

The informed consent form serves as an agreement between the investigator and the survey participant. It represents a pledge by the investigator as to what the survey participant can expect in terms of survey process, benefits, rights, risks, costs and potential harm. It must be written in simple, clear language that is native to or well understood by the participant, and be at a reading level that is appropriate for the population being surveyed. It should use active voice, short words and short sentences, and should not use coercive language.

During the informed consent process, potential participants must be given the following information:

- a description of the survey and the organization conducting it, the survey objectives, the populations being invited to participate and the number of expected participants;
- a clear indication that participation is voluntary and that participants have the right to withdraw from the survey at any time without penalty or loss of benefits;
- a description of the participant's role in the survey, including duration and frequency of participation, the procedures that the participant will undergo, the general nature of the interview questions, and the procedures for specimen testing;
- a description of the possible risks, burdens and discomforts of participation;
- a description of the direct benefits to the participant, or the indirect benefits to the community or society that are expected to come from the survey;
- measures made to protect the privacy and confidentiality of the participant and any information that the person provides;
- documentation that the protocol was reviewed and approved by an REC; and
- contact information and procedures for contacting the overseeing REC and principal investigator in case of further questions.

Informed consent can be given orally or can be written (e.g. initials, signature, mark or thumbprint). Particularly in surveys of stigmatized populations, investigators and participants may prefer to use oral informed consent because it does not produce a written record that could compromise participant anonymity. Some countries, however, require written informed consent. Anonymity and confidentiality can still be preserved in these cases. For example, participants can be asked to mark the informed consent form with an "X" to avoid providing any personal identifying information. If needed, the participant's unique survey identification number (ID) can be used to link the informed consent form with data-collection forms and biological specimens.

Voluntary agreement component of consent form

I understand what it means to join the survey. I understand my rights and risks. I had time to ask questions about the survey. I understand that I can join the survey at my free will. I understand that I can leave the survey at any time.

Have all your questions been answered?

☐ YES ☐ NO

Do you agree to do an interview?

☐ YES ☐ NO

Do you agree to the blood draw and testing?

☐ YES ☐ NO

Do you agree to the urine sample and testing?

☐ YES ☐ NO

Do you agree to the swab sample and testing?

☐ YES ☐ NO

Do you agree that we may keep the leftover blood/urine/swab for future testing?

☐ YES ☐ NO

If a national government requires that participants sign their name on the informed consent form, procedures must be put in place to ensure confidentiality of the information. This may include not linking informed consent forms to participant data and specimens, or using unique survey codes to link the forms, data and specimens; storing the informed consent forms securely and separately from data and biological specimens collected during the survey, as well as after the survey is complete; and minimizing access to the informed consent forms.

Participants must give informed consent for each element of the survey, including the interview, each specimen collected, and any storage or future testing that might be conducted on the stored specimens. Participants may agree to some procedures but refuse others. A single form can be used to obtain informed consent for all elements as long as the procedures to which a participant is consenting are clear (see the list of possible procedures in the box). Separate consent should be obtained for specimen storage and future unspecified testing. Since BBS are generally conducted anonymously, it is not typically possible to return results of future testing to participants. The consent form should indicate whether this is the case, and whether the participants should expect to benefit in some way from this testing.

Only adults are able to provide informed consent. The age of majority that defines adulthood should be considered during survey planning because it differs by country. Some individuals who have reached the age of majority are unable to provide consent; for example, adults with cognitive (mental) disabilities or people who are under the influence of drugs or alcohol. For these individuals, informed consent must be given by their legally authorized representative (e.g. parent or guardian), in accordance with national laws. And individuals unable to provide informed consent must still give assent, indicating their agreement to participate.

In cases where survey participants are too young to give their own informed consent but obtaining informed consent from their parent is inappropriate or could cause harm to the participant, investigators can seek a waiver for parental informed consent. An applicable situation would be a survey involving drug users aged under 18 years who are living and working away from their parents. The request for waiver must be submitted to, and reviewed and approved by, the governing REC or RECs for each protocol.

Although it is often advisable to provide participants with a copy of their consent form, investigators should use caution when doing so in surveys of stigmatized populations. Possession of the form might itself pose a threat to the participant's safety by revealing the person's risk group identity (e.g. MSM or PWID). Instead, investigators can provide participants with minimal information, such as the contact information of the investigators and appropriate RECs.

A-3.6 Ensuring privacy and confidentiality

Protection of a person's privacy and data confidentiality are essential for surveys. It is the investigator's responsibility to ensure that both are protected during and after the survey.

Privacy refers to the ability of an individual or group to control access to themselves or to information about themselves. The nature of privacy can vary, depending on the individual and the context in which the information is provided. For example, protecting a child's privacy may require the presence of the child's parent or legal guardian; whereas protecting a young person's privacy may be better achieved by withholding notice of survey participation from the person's parent or legal guardian. Aspects of privacy related to survey conduct include site selection and recruitment processes, the nature of the data collected, and data dissemination and use.

Site selection and recruitment

It is the investigator's responsibility to ensure the security of survey participants, particularly when the survey is conducted among vulnerable or marginalized populations. The investigator should choose a site and opening hours that are convenient for participants, and a site where they will feel safe from harassment by the general public or law enforcement. As far as possible, participants should not be visible to outsiders when participating in the survey. Responses to interview questions should not be audible to others at the survey site. The site itself and its location should be chosen to be discreet and to not draw unwanted attention, although whether this should be a busy location or a quiet place will depend on the context (e.g. the location could be near a frequented space such as a church, mosque or market, or on a quiet residential street with few pedestrians).

When selecting survey sites, investigators should consult with the target population and with individuals who provide services to them. For surveys involving populations that engage in illegal activities, investigators should seek guidance from the target population about whether to sensitize local law enforcement forces, to help minimize threats of harassment and raids during survey conduct.

Nature of data collected

To further respect the privacy of individuals, the survey should only seek to obtain and record information that is needed and will be used. For example, information about a participant's personal hobbies or home address are unrelated to the individual's HIV/sexually transmitted infection (STI) risk behaviours and should not be included in a survey.

Confidentiality

"*Confidentiality*" refers to the protection of the private information provided by the participant. During surveys, participants provide information with the expectation that it will only be disclosed according to processes outlined during informed consent. Survey staff should sign data confidentiality agreements that explain the procedures used to ensure data confidentiality.

Data dissemination and use

When disseminating results, care must be taken to ensure that the information does not reveal the identity of any individual participant or group of participants. For example, if the survey involved population mapping, the exact locations of hotspots should not be published because this might jeopardize participants' anonymity. Instead, these data could be aggregated at a neighbourhood or higher level, depending on the population or hotspot numbers.

A-3.7 Personal identifying information

"*Personal identifying information*" is information that can be used on its own or in combination with other information to reveal a person's identity. Among its many forms are names, phone numbers, physical addresses, email addresses, social security or other identifiable numbers, fingerprint or other images, and voice recordings.

Surveys should generally be conducted anonymously, meaning that personal identifiers should not be recorded on any forms, including informed consent forms, data-collection forms, biological specimens or referral forms. There should be no way to identify a participant from the survey data. Unique survey codes should be used in all records, and constructed without the use of personal identifying information such as complete date of birth.

A-3.8 Data security

Data collected during the survey should be managed in a way that protects the privacy of participants and the confidentiality of the information they provided. Investigators should develop clear data security and data-management in standard operating procedures (SOP), and should ensure that the SOP are understood by all involved in the survey, and are applied throughout data collection, transfer and storage. Access to collected data should be limited to staff on a need-to-know basis. Paper-based data forms should be stored in a locked room or file cabinet, and electronic data should be stored on password-protected computers or in password-protected files. Any data shared outside the survey team – such as with a third party conducting secondary analysis – should be anonymized.

A-3.9 Reimbursement to participants

Participants may be given reimbursement for travel costs, time spent, lost wages or other expenses incurred for participating in the survey. Additionally, they may be given free medical care. The reimbursement and medical care offered should not be so substantial that persons feel they need to participate. Investigators should be particularly sensitive to this situation in resource-limited settings, where even a small amount of reimbursement can have a major impact on a person's decision to participate in a survey. The type and quantity of reimbursement should be decided in consultation with the survey population and other stakeholders, and must be justified and approved by the governing RECs. In respondent-driven sampling (RDS) surveys, reimbursement is generally given both for survey participation and for the efforts of participants to recruit others for the survey.

A-3.10 Return of test results

In general, test results that may be of use for the participant's health care should be returned. Procedures for returning clinically meaningful test results to participants must be clearly described during informed consent. Confidentiality is best upheld using anonymous, linked specimen testing. For participants who wish to know the results of their HIV or STI tests, processes must be in place to allow easy return of test results and appropriate post-test counselling.

Where rapid testing is feasible, results can be returned during the initial survey contact. It is important to follow local testing algorithms for confirmation of test results, where needed, to comply with national policies. Where rapid testing is not available, it may be possible to have participants return to the survey site after a specified time frame in order to obtain their test results. In this case, it is important to implement measures to verify a participant's identity at the second visit. In some settings, it may be possible to return test results to a clinic or health facility of the participant's choosing.

A-3.11 Treatment and referral

Participants with test results indicating infection (e.g. HIV-seropositive) should be provided with information about appropriate treatment and care services. For conditions that can be treated safely by survey staff, same-day treatment should be provided (e.g. oral treatment with antibiotics for select STIs), in line with national policies and IRB approval. For infections that require complex or long-term treatment (e.g. HIV or viral hepatitis), efforts need to be made for a facilitated process of effective linkage to nearby health facilities, ideally to those with experience working with the survey population. Investigators should develop agreements with these service providers to ensure that referred survey participants are accepted. Investigators must ensure that the act of referral does not disclose information about the participant's risk group (e.g. MSM). For example, whereas it may be common to give participants a referral sheet with a list of clinics that have experience working with their population, such documents can indicate that a person participated in the survey but should not indicate which populations the survey targeted.

A-3.12 Adverse events, protocol deviations and protocol amendments

Adverse events

An adverse event is a negative event that happens to a participant during the conduct of the survey, regardless of whether it is related to survey participation. Adverse events may include unwarranted disclosure of a survey participant's identity, arrest or violence related to survey participation, errors in post-test counselling (communication of a wrong result), or allergic reaction to an antibiotic administered for a diagnosed STI. "Serious adverse events" are events that are life-threatening, require hospitalization, or result in death or persistent or significant disability. "Unexpected adverse events" are those that were not cited by the investigator in the protocol as a possible consequence of participation.

Reporting of adverse events

Investigators should promptly report unexpected adverse events to the governing REC according to the REC's guidelines. Often, the REC will require an initial informal notification via email within a few days. Subsequently, the REC may expect a formal incident report within a given time, often 1 week. The investigator should develop or use the REC's adverse event forms to document and report such events, including a description of the adverse event, the date it occurred, responses taken to the event, and the date it was reported to the governing REC.

Protocol deviations

Protocol deviations are unintended survey procedures or events that were not described in, or that differ from, the approved protocol (e.g. redefinition of the sampling domain, or survey participation by an individual that did not meet eligibility criteria). Protocol deviations may or may not put participants at risk. They may even be necessary in emergency situations. However, all protocol deviations should be reported to the overseeing REC according to its guidelines, and may ultimately require that the protocol be amended.

Protocol amendments

All approved survey materials must be used as approved by the governing REC. This applies not only to the protocol but also to informed consent documents, data-collection tools and recruitment materials. If investigators wish to alter some of the procedures described in the approved protocol, a protocol amendment must be prepared and submitted to the REC for review and approval. The proposed changes can take effect only after approval is given.

A-3.13 Research ethics training

Investigators and survey staff interacting with participants should receive training in human subjects research ethics. The curricula for such training generally cover ethical principles of research, REC roles and responsibilities, participant privacy and confidentiality, voluntary informed consent procedures, and reporting of adverse events and deviations from the protocol. Examples of relevant courses include the Collaborative Institutional Training Initiative course (United States of America, USA), the National Institutes of Health's Protecting Human Research Participants course (USA), Training and Resources in Research Ethics Evaluation (Switzerland, available in English, French, German and Portuguese), the Epigeum course (United Kingdom), the Fundacion Bioetica course (Spain) and FHI 360's Research Ethics Training Curriculum (USA).

A-3.14 Special considerations for incarcerated populations

Incarcerated populations are especially vulnerable; hence, special considerations must be taken into account when surveying this population. Incarcerated populations tend to be of lower socioeconomic status, have lower education attainment and reading level, and have higher rates of infectious disease, chronic illness and mental illness than the general population. They may also have higher rates of substance and alcohol use, and be more prone to injury, violence and abuse. Within a prison setting, inmates generally have poorer access to health and legal services, and little or no personal privacy. Two major considerations for surveys among incarcerated populations are the ability of prisoners to provide informed consent and their right to privacy, confidentiality and autonomy.

Right to privacy, confidentiality and autonomy

Maintaining privacy is an enormous challenge in prison settings, because inmate movements are closely monitored by both prison officials and other inmates. Providing health information and services to a prisoner in a confidential setting may be equally challenging. In addition, prisoners may fear repercussions from prison officials or other inmates, including threats, physical or sexual abuse, punishment or lengthened prison terms, as a result of their decision to participate or not. Special care must be taken in designing the survey's methods for recruitment, testing and provision of test results, and follow-up care (treatment and referrals) to ensure that potential participants' right to privacy, confidentiality and autonomy is upheld throughout the process. It is the investigator's responsibility to consult with local experts to ensure that the methods used are tailored and acceptable to the local context.

Prisoners may also feel coerced into participating by prison authorities or compelled to participate in order to access the health services offered as part of the survey. This is particularly relevant for surveys conducting HIV and STI testing in prisons, where routine access to such testing is often limited or nonexistent. Where possible, investigators should seek to minimize the discrepancy in health services offered to participants and nonparticipants. For example, confidential HIV and STI testing services could be made available to all inmates during the time of the survey, but only data from those selected to participate would be recorded and reported by the survey team.

Ability to provide informed consent

Incarcerated populations may read at lower levels. Therefore, investigators must adjust the language used in informed consent forms to ensure that participants are able to fully understand the survey's purpose, risks and benefits, and procedures, and the steps taken to ensure privacy and confidentiality.

A-3.15 Web-based surveys

Investigators increasingly use the internet to conduct surveys, and current guidelines may not address all of the issues raised by the unique characteristics of an internet survey. In addition to the ethical considerations applicable to conventional surveys, investigators must consider the following aspects:

- IRBs may regard internet protocol (IP) addresses as personal identifiers. Surveys that strive to maintain anonymity need to take measures to prevent the unintended collection of IP addresses, or else take measures to protect the confidentiality of the individuals associated with these IP addresses.
- internet-based surveys may face particular challenges to assess eligibility criteria; for example, it may be difficult to assess age, and this could lead to the enrolment of minors in a survey designed to be limited to adults.

A-3.16 References

1 The Nuremberg Code. 1947 (https://history.nih.gov/research/downloads/nuremberg. pdf, accessed 25 August 2016).

2 WMA. Declaration of Helsinki – ethical principles for medical research involving human subjects (2013 update). World Medical Association (WMA); 1964 (https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ accessed 22 May 2017.

3 The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: ethical principles and guidelines for the protection of human subjects of research 1979 (http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html, accessed 25 August 2016).

4 CIOMS/WHO. International ethical guidelines for biomedical research involving human subjects. Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with World Health Organization. Geneva, Switzerland: CIOMS, Council for International Organizations of Medical Sciences (CIOMS)/ World Health Organization (WHO); 2002 (http://cioms.ch/publications/layout_guide2002.pdf, accessed 25 August 2016).

5 WHO/UNAIDS. Guiding principles on ethical issues in HIV surveillance. Geneva: World Health Organization (WHO)/Joint United Nations Programme on HIV/AIDS (UNAIDS); 2013 (http://www.who.int/hiv/pub/surveillance/2013package/module2/en/, accessed 3 August 2016).

4. Formative assessment

This chapter describes formative assessment, one of the most important parts of the survey process. A good formative assessment will inform the design and implementation of a biobehavioural survey (BBS), including how participants are recruited and operational processes are carried out.

The assessment also provides an opportunity for investigators to develop relationships with the target population, build trust and solicit their input on how the survey should be carried out. A BBS that is informed by the findings of a formative assessment is often more acceptable to the target population, more likely to succeed and less likely to result in unintentional harm to the target population than a survey that has not been subject to such an assessment.

A-4.1 The purpose of formative assessment

A formative assessment is used to collect initial information about a target population, gain that population's support for a BBS, and inform the development and conduct of the survey itself. Formative assessment is most important in settings where little is known about the target population. However, even when there is abundant knowledge about the target population, a formative assessment can help to re-engage them, identify changes since the last survey and learn from the population's experiences with previous surveys. In all settings, formative assessment can provide immediate feedback on service availability.

Formative assessment should be conducted after the target population has been selected, and in every location and for every population in which a BBS is planned. Qualitative methods are most commonly used for formative assessment data collection, but quantitative methods may also be used.

The results of the formative assessment may shape the design of the BBS, particularly in settings where the investigators have little or no pre-existing relationship with the target population. It is thus good practice to develop separate protocols for the formative assessment and the actual BBS, so that findings from the formative assessment can be used to develop or modify the BBS protocol. For example, if the formative assessment shows that the target population is socially well connected but does not congregate at specific locations, investigators may choose to use respondent-driven sampling (RDS). Conversely, if the formative assessment shows that the population is not socially well connected but does congregate at specific locations, the investigators may choose to use time-location sampling (TLS).

Investigators should consult with members of the target population and meaningfully engage them in all stages of the BBS process, beginning with the formative assessment. It is also good to include community-based organizations – such as those that serve or are composed of people who inject drugs (PWID) or sex workers (SW) – in survey design and implementation. Often, trust that investigators build with target population members during the formative assessment will result in greater willingness on the part of the population to connect investigators with other key informants and venues where the target population

Key terms

Formative assessment:	The collection of initial information about a population to inform the preparation and conduct of a survey of that population.
Free listing:	Asking respondents to list as many items as will fit in a particular domain (e.g. all the different drugs that are taken by local drug users).
Pile sorting:	Asking the respondent to group items from the same domain, often using cards; for example, sorting the names of the locally used drugs that are either stimulants or depressants.
Venue:	In terms of BBS, a place where the target population congregates (e.g. bars or brothels).

congregates. The formative assessment is also a good opportunity to identify potential survey staff, venues for TLS, or seeds for a BBS using RDS.[1]

A-4.2 Formative assessment objectives

This subsection describes the objectives of formative assessment, which are to:

- understand the target population and context;
- identify existing services and gaps;
- inform survey methods (e.g. sampling strategy, questionnaire and biological specimen collection); and
- engage stakeholders.

These objectives are discussed below.

A-4.2.1 Understand the target population and context

A BBS is more likely to succeed if investigators have a collaborative and respectful relationship with members of the target population. The relationship between investigators and the target population may take time to develop, and may be threatened if the expectations of the investigators and the target population are far apart. It is important to spend sufficient time and effort to reach a common understanding with the target population about key issues of the survey, and to maintain this collaboration throughout the survey until final results are disseminated.

During the formative assessment, investigators should learn about the social and legal environment of the target population, the language and questions that are most appropriate for use in the survey questionnaire, and the health service needs of the target population. Formative assessment should answer the following questions:

- who is the target population?
- what is the legal and social context of the target population, and how has it changed since the last survey?
- where can investigators locate members of the target population and engage them in the survey? How do members of the target population interact with one another?
- was the target sample size achievable?
- What are the sexual and drug-taking practices of the target population, and how do they talk about them?

Characterize the target population

Before conducting a BBS, investigators should be able to roughly describe the target population, including the sex, age range, nationality, marital status, ethnicity, neighbourhood of residence, employment and subgroups. For RDS surveys, this information will assist investigators in selecting diverse seeds. And for all sampling methods, this information can later be compared to the BBS sample to identify subpopulations that are under-sampled or excluded.

Understand the legal and social context of the target population, and how has it changed since the last survey

Investigators should be able to place the target population in their social context based, for example, on their demographic characteristics, risk behaviours and laws that affect them. For instance, investigators might observe that PWID have started using new injection methods, or that their social visibility is increasing or decreasing. Legal prohibitions on commercial sex, homosexuality, or injecting drug use can discourage participation in the survey, affect the openness with which the team conducts the survey and have ethical implications. Accordingly, the legal context must be well understood so that investigators can ensure the safety and security of their staff and participants.

Identify where investigators locate members of the target population and engage them in the survey and in understanding how members of the target population interact with one another

To choose the most appropriate sampling methodology, investigators must understand where and how members of the target population interact with one another, including interaction between subgroups. For example, knowing that SW are largely based in brothels is not sufficient for determining whether it will be more appropriate to recruit them through TLS. Brothel-based SW can be very different from street-based SW, and cluster sampling methods may yield a sample that does not represent both populations. Similar challenges may exist in circumstances where the target population is ethnically diverse, or includes both native and foreign members but the sampling frame captures only one of these groups.

Determine whether target sample size is achievable

During the formative assessment, investigators should solicit rough estimates of the size of the target population or the share of the total population they represent. If the target population is too small to obtain a representative sample, investigators may opt to conduct a nonrepresentative BBS or a qualitative survey of risk behaviours and service needs. For RDS

[1] In RDS surveys, "seeds" are the individuals who are chosen by the investigators and represent the start of the recruitment chains.

surveys, investigators should inquire about participants' social network sizes, to determine whether they are large enough for peer-referral sampling. Methods for estimating population size for programme planning, monitoring and epidemic modelling are described in Section A-11.

Understand the sexual and drug-taking practices of the target population, and how the population talks about those practices

During the formative assessment, investigators should learn about the sexual and drug-taking behaviours of the target population, and how they talk about those relationships and practices. The information gained will allow investigators to adapt standard data-collection tools, such as those presented here, to the local context.

BBS results are only valid if the participants accurately understand and respond to the survey questions. Participants and investigators may use terms differently. For example, participants may use the term "bisexual" to describe having had more than one sexual partner, whereas investigators may use it to describe someone who identifies as being sexually attracted to men and women. Similarly, individuals with "sugar daddies" (i.e. sexual partners who regularly give them money or gifts) may not perceive this as a "transactional" sexual relationship and may use another term for this activity, or not label the activity with any term at all. Other categorizations – for example, "regular", "casual" and "commercial" – may mean little to them. Similarly, target populations may use specific local terminology to describe sexual or drug-related behaviours, or may redefine terms (e.g. using "shooting up" to indicate injecting drugs).

A-4.2.2 Identify existing services and gaps

Formative assessment findings about service accessibility can prompt action long before final results from the BBS are available. The formative assessment can yield information about the target population's health needs, the services provided to them, and their access to those services. Service providers can respond to gaps in service; for example, by improving access to HIV testing, condoms and lubricants. Later, final results from the BBS can be used to fine-tune or expand services. An understanding of the available services is also needed for efficient referral of participants during the survey.

A-4.2.3 Inform survey methods (e.g. sampling strategy, questionnaire and biological specimen collection)

As noted above, the formative assessment should inform the design of the BBS – specifically the sampling strategy and survey methods – and mobilize and engage the target population. The formative assessment can help investigators to understand where the target population is located, how members interact, whether they are willing to recruit one another into a survey and whether there is another method for identifying population members (e.g. a list kept by service providers). This will help investigators to select the optimal sampling strategy. For more on sampling strategies, see Section A-9.1.

The formative assessment can also help gauge the target population's willingness to provide biological specimens for HIV and sexually transmitted infection (STI) testing, the preferred location and timing of the survey, characteristics of survey staff, coupon design for RDS surveys, the desired method of reimbursement for participation, and whether authorities should be involved in or informed of the survey. The assessment may also yield seeds for an RDS survey, venues for TLS, and an understanding of local terminology; for example, terminology about various sex partner types or how population members differentiate between paying and nonpaying partners.

The formative assessment can also reveal barriers to survey participation and can help to improve the accuracy with which the results are interpreted. For example, whereas BBS data may suggest that older men who have sex with men (MSM) make up a minority of the target population, formative assessment results may show that older MSM are really just less willing to participate in the survey than younger MSM. Similarly, the formative assessment may help to explain why SW who operate on the street participate at a different rate than those in brothels or guesthouses.

Finally, the formative assessment is an opportunity for investigators to learn the terminology that the target population uses to describe the behaviours they engage in. This will allow investigators to design a survey questionnaire that is relevant to and easily understood by the target population.

A-4.2.4 Engage stakeholders

Engaging a broad group of stakeholders will improve the acceptability, quality and credibility of survey results. It may also help to ensure that the survey does not duplicate research already undertaken. Also, such broad engagement will build national and local interest from all sectors of society, particularly the target population, so that all partners will be willing to act on the findings of the BBS. Investigators should seek to engage stakeholders throughout the entire survey process, from the formative assessment to the BBS, and finally to the data-use stage.

A-4.3 Conducting a formative assessment

The structure and extensiveness of the formative assessment will depend in part on how much is already known about the target population, and on the strength of the existing relationship between the investigators and the target population.

The first step in the formative process is usually a review of existing publications about the target population, including those describing their sociocultural history and their epidemiological characteristics. Investigators should review peer-reviewed manuscripts and the "grey" literature (e.g. surveillance reports, programme reports and conference abstracts). Where little or no country-specific data are available, investigators may review data on target populations from the same region.

Over the past decade, there has been an increasing appreciation for qualitative methods in the formative assessment phase of a BBS *(1, 2)*. Two methods are now typically used: key-informant interviews and focus groups. Other participatory methods include free listing and pile sorting – which may help to establish appropriate language for use in survey questions, informed consent forms and transition text in the surveys – and mapping and ranking *(3)*. More information on qualitative methods, including interview guide development and data analysis, can be found in Chapter A-12.

A-4.4 Optimal use of information gathered through the formative assessment

Data from the formative assessment can inform all aspects of BBS protocol development. The findings of the BBS should also be reviewed in light of the formative assessment results, to ensure that the results are interpreted accurately. Formative data may also help to improve service provision and policy-making. For example, service providers may believe that they are providing services to MSM; however, if MSM cannot name a single organization that targets them, that disconnect is important to recognize, and points to an opportunity for MSM to be further engaged in identifying the necessary services and their delivery.

Investigators should start working on the BBS protocol as soon as the formative assessment has helped to identify the most appropriate recruitment method. Meanwhile, formative assessment findings can be disseminated through a report, and investigators should present them to key stakeholders.

A-4.5 Ongoing formative assessment

Ideally, formative assessment is an ongoing activity that should continue during the BBS. Investigators should regularly consult the target population in order to identify issues that might affect participation or survey conduct. For example, police may have begun targeting the population with arrests, or reimbursement for participation may be too low to engage a sufficient sample size; conversely, reimbursement may be too high, encouraging ineligible individuals to fake membership in the target population in order to participate. This consultation can include conducting exit interviews of survey participants, having an anonymous suggestion box for participants to leave comments, or discussions with stakeholders, including community-based organization staff members who work with the target population. Exit interviews and the questions that will be asked should be included in the formative assessment protocol. Findings from ongoing formative assessment can be used to revise and strengthen survey procedures.

A-4.6 References

1 Johnston LG, Whitehead S, Simic-Lawson M, Kendall C. Formative research to optimize respondent-driven sampling surveys among hard-to-reach populations in HIV behavioral and biological surveillance: lessons learned from four case studies. AIDS Care. 2010;22(6):784–792.

2 Simic M, Johnston LG, Platt L, Baros S, Andjelkovic V, Novotny T et al. Exploring barriers to 'respondent-driven sampling' in sex worker and drug-injecting sex worker populations in Eastern Europe. J Urban Health. 2006;83(6):i6–i15.

3 International HIV AIDS Alliance. Tools together now! 100 participatory tools to mobilize communities for HIV/AIDS. Brighton: International HIV AIDS Alliance; 2006 (http://www.aidsalliance.org/assets/000/000/370/229-Tools-together-now_original.pdf?1405520036, accessed 4 August 2016).

5. Questionnaire development

This chapter provides an overview of various methods for collecting survey data, and for developing and testing questionnaires before implementation.

The questionnaire module section (Section II) of these guidelines contains standardized instruments for collecting survey data among key populations – men who have sex with men (MSM), people who inject drugs (PWID), sex workers (SW) and transgender persons (TG) – and other vulnerable populations. These instruments have been field tested and, where indicated, the modules use validated questions and scores. Investigators are strongly encouraged to use these "established" questionnaires whenever possible, for two main reasons. First, questionnaire development is a difficult process; it requires extensive preparatory work, including cognitive testing and evaluation. The instruments provided use questions, time references and skip patterns that have been tested and are known to produce high-quality data. Second, these instruments include questions that have been used throughout the world; thus, their continued use will enable researchers to compare survey results across countries.

The standardized instruments provided with these guidelines were designed to be self-administered using electronic data-collection methods, which may complement or replace paper-based data collection. These instruments may be used in face-to-face or computer-assisted self-interviews, will work on multiple electronic platforms (e.g. tablets, laptops or smartphones), may be adapted to a variety of local contexts and may be modified for application using paper-based data-collection methods.

When planning a survey, investigators should be familiar with the advantages and disadvantages of various data-collection methods in order to choose the method most appropriate for the circumstance.

A-5.1 Key steps in questionnaire development

The key steps in questionnaire development, described in detail below, are as follows:

Key terms

Cognitive:	Relating to thought processes and gaining knowledge (through activities such as thinking, understanding, learning and remembering) by the use of reasoning, intuition or perception.
Cognitive testing:	Method to investigate how well a question performs and whether respondents understand the question correctly.
Distal determinant:	With regard to HIV, a structural (not biological or behavioural) variable that affects HIV risk (e.g. stigma and discrimination, community mobilization, violence and poverty).
Proximal determinant:	With regard to HIV, a biological or behavioural variable that directly facilitates HIV acquisition or transmission.
Recall bias:	Systematic error due to differences in accuracy or completeness of remembered past events.
Saliency:	Degree to which a topic or event resonates with a potential survey participant.
Salient:	Notable or important.
Syndemic	Two or more endemics or epidemics occurring simultaneously and acting synergistically.

1. decide on methods for questionnaire (interview) administration.
2. determine investigation topics (questionnaire domains).
3. develop and adapt the questionnaire.
4. translate and back-translate the questionnaire.
5. conduct cognitive testing.
6. pretest the questionnaire.
7. train interviewers.
8. pilot test the survey tool.

A-5.1.1 Decide on methods for questionnaire (interview) administration

Questionnaire delivery mode

The questionnaire may be delivered by an interviewer or may be self-administered by the participants themselves.

Some standardized questionnaires are designed for a specific delivery mode, whereas others may need to be modified. When deciding on the method of delivery, investigators should consider factors such as the length and complexity of the questionnaire, the literacy of a typical participant and the sensitivity of the questions. There are trade-offs to each of these factors. Formative assessment and, if applicable, pilot testing of different interview administration methods may help investigators to make an informed decision well in advance of survey implementation, thus allowing time for modification.

Interview questions and instructions should be clear. For self-administered interviews, questions and instructions need to leave no room for misunderstanding because participants may not feel comfortable asking for clarification. Misunderstanding of the questions may result in invalid answers.

Data collection mode

There are two data collection modes – paper based and electronic – as discussed below.

Paper-based data collection

Paper-based questionnaires may be delivered by an interviewer or may be self-administered. When using paper-based questionnaires to collect data, additional time should be allotted for training interviewers. Also, the interviewers should have previous experience collecting data from complex surveys on paper.

To complete self-administered paper-based questionnaires, participants must be moderately literate and able to read, understand and follow instructions for completing the form. The number and complexity of "skip" instructions should be limited, especially for self-administered questionnaires. Data may be entered electronically either shortly after the interview or later in batches. Double data entry is strongly recommended to limit data entry errors.

Electronic data collection

Electronic data collection instruments can also be delivered by an interviewer or self-administered. They may be developed using a variety of existing software; for example, the Questionnaire Development System (QDS™ Nova Research, Bethesda, MD) and Open Data Kit (ODK). Table A-5.1 summarizes various systems of electronic data collection, and their advantages and disadvantages.

Table A-5.1 Systems of electronic data collection, and advantages and disadvantages

System	Mode of delivery	Advantages	Disadvantages
Computer-assisted self-interview (CASI)	• Participants self-administer the questionnaire	• Allows participants to answer sensitive personal questions privately on a computer • Eliminates the need for subsequent data entry and the risk of introducing errors • Ensures questionnaire is delivered consistently – participants all experience the questions in the exact same way • Flexible and convenient: - can be completed with a mouse or a touch screen - can be segmented so that some sections are self-administered and some are administered by an interviewer - handles skip patterns automatically - can be offered in multiple languages - works for single site or multisite studies - data are readily exported for analysis	• Requires participants to have a certain degree of literacy • May be less suitable for participants who are unfamiliar with or uncomfortable using a computer • May not offer sufficient opportunities for participants to seek clarification about individual questions, resulting in misinterpretation of their meaning and invalid responses
Audio computer-assisted self-interview (ACASI)	• Same as CASI but with the addition of an audio component, so participants can listen to questions (e.g. through a headset) as well as read them	• More appropriate than CASI for collecting data from participants with limited reading ability; otherwise, same advantages as CASI	• Same disadvantages as CASI

System	Mode of delivery	Advantages	Disadvantages
Computer-assisted personal interview (CAPI)	• A face-to-face interview in which the interviewer immediately enters the response data into the computer	• Eliminates the need for subsequent data entry and the risk of introducing errors • Ensures questionnaire is delivered consistently – participants all experience the questions in the exact same way • Interviewer is immediately available to clarify questions and responses • Flexible and convenient: - works with either a mouse or a touch screen. - can be segmented so that some sections are interviewer-administered and some are administered in private (via CASI or ACASI) - handles skip patterns automatically - can be offered in multiple languages - works for single site or multisite studies - data are readily exported for analysis	• May result in response bias if participants are not comfortable being honest with the interviewer
Handheld-assisted personal interview (HAPI)	• Similar to CAPI, a face-to-face interview in which the interviewer immediately enters the response data into a handheld computer (e.g. personal digital assistant (PDA) or mobile telephone)	• Same advantages as CAPI	• Same disadvantages as CAPI
Web-based	• A questionnaire is displayed on a website and participants log on and enter their answers[a b]	• Allows participants to answer sensitive personal questions privately on a computer • Participants can complete the questionnaire on their own schedule and over several sessions, if necessary • Handles skip patterns automatically	• Requires participants to have a certain degree of literacy • May be less suitable for participants who are unfamiliar with or uncomfortable using a computer or the internet, and inappropriate for populations entirely without access to them • The clarity of the questions and instructions are especially important as participants may have even less opportunity to seek clarification from survey investigators • Potential participants must be given the web address and invited to complete the survey through some type of contact, email list, or advertisement • Requires a website host that is always accessible and fast enough to handle multiple participants simultaneously

[a] For online survey, see http://www.orau.gov/cdcynergy/soc2web/Content/activeinformation/tools/toolscontent/surveyselfadministered_internet.htm (accessed 17 June 2014)

[b] Dillman 2000 (1)

Several commercial vendors offer web-based survey development services, including Survey Monkey[1] and GfK Knowledge Networks.[2]

[1] http://www.SurveyMonkey.com

[2] http://www.knowledgenetworks.com/ganp/

Proximal and distal determinants of HIV

The spread of HIV is due to both "proximal" and "distal" or "structural" determinants *(2-4)*. Proximal determinants of HIV are biological or behavioural variables that directly facilitate the acquisition or transmission of HIV. Biological variables include HIV viral load and subtype, and the presence of other infections such as sexually transmitted infections (STIs). Behavioural variables include the type (e.g. vaginal or anal) and frequency of sex, condom use, and the types and number of sex partners.

Distal or structural determinants can be just as powerful as proximal variables, but they affect HIV risk through the proximal variables. These determinants include factors such as stigma and discrimination, community mobilization, violence and poverty. The prevalence and relationship of distal determinants to risk behaviours, and ultimately to HIV prevalence and viral load suppression, are important for informing HIV interventions *(5-7)*.

An emerging recognition of "syndemics" has helped to show that HIV epidemics do not occur in isolation. Rather, they often occur in tandem with other epidemics and need to be addressed comprehensively *(8)*.[i] The best known syndemics co-occurring with HIV are STI, tuberculosis (TB) and injecting drug use. Equally important syndemics include psychosocial issues, alcohol use and violence *(9, 10)*.

A-5.1.2 Determine investigation topics (questionnaire domains)

Investigation topics

All biobehavioural surveys (BBS) should collect standard information on demographics, risk behaviours and access to HIV testing, prevention and health-care services. BBS data-collection instruments may include questions to assess both proximal and distal determinants of HIV.

Questionnaire domains may include but are not limited to:

- demographics
- sexual and reproductive health
- sexual behaviour
- alcohol and drug use
- HIV services uptake
  - access to counselling and testing
  - knowledge of serostatus
  - access to care and treatment
  - retention in care
- access to other health-care services
  - general
  - STI
  - tuberculosis (TB) (HIV coinfected)
- mental health, including depression
- psychosocial support
- shame, stigma and discrimination in the community and health facility and violence
- knowledge of HIV prevention and transmission

Additional domains may be added depending on the target population, survey context, and information needed to understand and respond to the HIV epidemic.

A-5.1.3 Develop and adapt the questionnaire

General considerations

When developing or adapting the questionnaire and selecting the method of administration, investigators must consider the points listed below:

- choice of variables – The data obtained should be those that are necessary to better understand the target population and respond to the epidemic.
- recall period – Participants are typically asked to recall incidents or behaviours occurring within a specified time period. Shortening this time period may reduce recall bias (i.e. systematic error due to differences in accuracy or completeness of remembering events) but may also limit the ability to measure the specified behaviour, because some participants will not have engaged in a particular behaviour during the shorter time period.
- order of questions – Questionnaires should open with easy, salient and nonthreatening questions (*11*). The interview should move from more general questions to more specific ones, and questions that reference dates should be asked in consistent chronological order. When changing topics, the interviewer should use transitional phrases to allow participants to adjust their thinking.
- length of the interview – The questionnaire should be kept as short as possible by removing questions that are redundant, yield the same answer from all participants, or are unlikely to be analysed or used (*12*). The saliency of the topic to the survey participant (see text box) will be a major determinant of the length of the questionnaire. Individuals on HIV treatment have more questions to answer than those who are HIV negative. Longer interviews may result in higher

survey refusal rates or participant fatigue, and unhappy participants may discourage others from participating. Survey investigators should pretest the questionnaire to determine how much time participants are willing to devote to the interview and adjust the length accordingly.

- validation rules – To minimize errors in recording data, investigators should build validation rules into the questionnaire for continuous variables (i.e. responses that take the form of numbers, including age or number of sex partners). For electronic interviews these rules (e.g. lower or upper limits) can be programmed into the software so that it automatically detects invalid responses. Validation rules can also be used with paper-based data collection. This requires substantial interviewer training to detect invalid responses.
- mandatory fields – It is almost always better to record a response such as "Don't know", "Refuse to answer" or "Not applicable" than to leave a response field blank. Blank response fields may be misinterpreted as missing data.
- comment and open-text fields – Occasionally, it is necessary to ask an open-ended question and record the response as text; for example, in the field "Specify if other". But open text can be difficult and time consuming to clean and categorize, so these fields should be used sparingly if at all, especially if they have little analytical value.

Adapting the data-collection instrument for local use

The standardized data-collection instrument in these guidelines comprises several questionnaire modules. Each module covers a specific topic and includes questions deemed important for better understanding the epidemic or for satisfying reporting requirements.

Within each module, questions are designated as "core" if the information being collected is considered essential (e.g. standard demographic information or information required to characterize risk) or actionable (e.g. able to guide public health action), and "secondary" if the information is useful but not essential. Questions that are core for some target populations may be secondary or not applicable for others. Survey investigators may build their questionnaire by choosing modules that are relevant to their country and the target population. This modular approach allows investigators to include additional topic areas by adding locally relevant modules (e.g. hepatitis and reproductive history) as appropriate, or omit modules that are irrelevant.

All questions and instructions should be designed so that they may be self-administered using computer-assisted self-interview (CASI) or audio computer-assisted self-interview (ACASI). Respondents can read questions, and with ACASI also listen to them through a headset, and choose response options. Alternatively, survey investigators may opt to use the questionnaires for face-to-face interviews.

The questionnaire modules included with these guidelines are listed on the next page. The description and rationale for each questionnaire module is provided in the introduction to the module.

Saliency

Saliency is the degree to which a topic or event is relevant to a survey participant *(13)*. Topics of great importance to the participant are said to have a high degree of saliency. Saliency operates at two levels, that of the question and that of the survey as a whole. In regards to individual questions, saliency refers to the importance of an event or action in a person's life. Important events are better remembered than those with less saliency. For the survey as a whole, saliency refers to the degree to which the subject matter of the survey is relevant to the target population. If the questions being asked are of great interest to the typical member of the target population, the survey is said to be highly salient.

Saliency has implications for survey investigators. Gaining cooperation or attaining a high response rate is more difficult when the saliency is low, because the target population has little motivation to participate. It is much easier to achieve a high response rate with a questionnaire that is highly salient and has a low participant burden (i.e. takes little time to complete, and is easy and straightforward to understand) than with one that has low saliency and a high participant burden. In the latter case, it may be necessary to give participants more reimbursement for their time in order to improve the response rate.

Questionnaire Modules

1. Questionnaire parameters
2. Eligibility
3. RDS recruiter–recruit relationship
4. Respondent-driven sampling (RDS) personal network size
5. Size estimation
6. ACASI tutorial
7. Demographics
8. Transgender characteristics
9. Sex work characteristics
10. Clients of sex workers characteristics
11. Incarceration characteristics
12. Transport work characteristics
13. Reproductive health
14. Sexual history
15. Recall sexual behaviour
16. Partner concurrency
17. Last sex act
18. Condom use/accessibility
19. Lubricant use
20. Microbicide use
21. Alcohol use
22. Noninjecting drug use
23. Injecting drug use
24. Drug overdose
25. Sexually transmitted infections
26. Hepatitis B virus/hepatitis C virus (HBV/HCV)
27. Services uptake
28. Post-exposure prophylaxis/pre-exposure prophylaxis (PEP/PrEP)
29. Shame, stigma, harassment and discrimination
30. Physical violence
31. Sexual violence
32. Internalized homophobia
33. Depression
34. Suicide ideation and attempts
35. Social support
36. Social cohesion, social participation and collective agency
37. Game of contacts
38. HIV knowledge and perceptions
39. Questionnaire feedback/interview status
40. RDS peer recruitment

Data measures

Survey investigators should review existing data reporting requirements and indicators such as those from the Joint United Nations Programme on HIV/AIDS (UNAIDS); the United States President's Emergency Plan for AIDS Relief (PEPFAR); and the Global Fund to Fight AIDS, Tuberculosis and Malaria, to:

- determine whether the questionnaire collects the information needed;
- identify gaps and determine whether new questions are needed; and
- consider how the data will be tabulated or analysed.

It is important to know which denominators will be used for each variable. For example, the denominator used to calculate the share of the population that used a condom at last sex with a casual partner should include only those who reported having sex with a casual partner.

Accordingly, the survey must include a question that asks whether the participant has had sex with a casual partner. Substituting in the number of participants who reported having sex with any type of partner could result in an underestimate of the indicator. For more information see Section 3, which provides an overview of data measures and suggests how they should be used. That section includes reference sheets for a core set of data measures considered critical for understanding the context in which key populations live and make decisions, and for monitoring and evaluating programmatic responses.

A-5.1.4 Translate and back-translate the questionnaire

Many surveys will use more than one interview language. Translators should have knowledge of health terminology, colloquial phrases and jargon used by the survey population. When translating questionnaires into other languages, it is recommended to back translate the questionnaire into the original language. Staff not involved with the translation should do the back translation without using the original-language questionnaire. The two questionnaires (i.e. original and back-translated) should then be compared, and any question and answer phrasings that differ in their meaning should be noted. The translation should be changed as necessary until the two questionnaires fully agree in their question and answer phrasing.

A-5.1.5 Conduct cognitive testing

Often, survey investigators want to add a few questions to the standardized questionnaire. Also, in some cases, a country may want to collect information for which no standardized questions exist. New, previously unused questions for a country should be evaluated for validity and to determine what information they are actually capturing. The goals of evaluation are to assess how respondents interpret survey questions and identify potential response problems that could introduce errors in the survey data. Questions should be designed based on the results of this evaluation, not on expert opinion *(14)*.

After a formative assessment has been conducted to identify topics of inclusion in the questionnaire, questions should be cognitively tested to ensure that they are properly understood *(15)*. The primary purpose of cognitive testing is to investigate how well questions

perform when asked of survey participants; that is, to test whether participants understand the question correctly and can provide accurate answers *(16)*. Cognitive testing ensures that a survey question captures the investigators' intent and makes sense to participants.

Cognitive testing includes the following steps:

- administer draft questionnaires and collect additional information about responses:
  - facilitate participants to share their thoughts while hearing and responding to questions;
  - identify specific problems and answer specific questions;
- use information to:
  - evaluate the quality of the question;
  - understand whether the question gathers the intended information;
- look at question-answering from the participant's perspective in order to understand:
  - the thought processes used to answer questions;
  - how questions are interpreted; and
  - how participants understand concepts.

Cognitive testing can help to ensure that survey questions capture the intended information. If they do not, investigators can modify questions well before they start the survey. Investigators should identify and document what each question should measure, and compare it to what it does measure. For example, for investigators, the term "sex" may refer to only vaginal and anal sex, but for participants it may also include oral sex. This process may thus identify subtle but important differences in interpretation. It should be well documented and capable of replication.

Cognitive testing should answer the following questions:

- how do respondents interpret the survey question?
- do respondents interpret questions differently than intended?
- how do respondents determine their response?
- does the question have the same meaning in all the languages in which it is asked?
- does the question collect the same type of information across cultures, languages, subgroups or settings (e.g. urban and rural)?

A-5.1.6 Pretest the questionnaire

Whether survey questionnaires are self-administered or implemented through face-to-face interviews, data-collection instruments must first be pretested. Pretesting is different from cognitive testing; its purpose is for trained interviewers to practise using the instruments as a means for improving data collection. Pretesting helps survey investigators to identify and address problems before they occur in actual survey implementation. Pretesting can also identify problems introduced by translation or the adaptation of an existing questionnaire to the local context. Survey staff should pretest all data-collection forms with diverse members of the target population, ideally under typical field conditions. Pretesting activities and findings should be documented to inform the revision of data-collection tools.

A-5.1.7 Train interviewers

If the survey is to use face-to-face interviews, investigators should develop an interview guide for the interviewers and their supervisors. The interview guide should address each interview question, explaining its rationale and its intended meaning, skip patterns, legal value definitions, appropriate background computation (if applicable) and informational messages. This will be helpful for training and to clarify any ambiguities or misunderstandings that may arise during implementation.

Interviewers should be thoroughly trained to ensure that the interview process is as standardized as possible. All interviewers should administer the questionnaire or other data-collection instruments in the same way to avoid introducing bias in how participants respond. This standardization should extend to how the questions are read and explained, and to the attitude the interviewers present to the participants. This can greatly influence the given responses – especially if the interviewer asks about illegal or stigmatized behaviour. To increase the likelihood of obtaining honest responses, interviewers should assure participants of their confidentiality, and should be thoroughly trained in open and nonjudgemental interview techniques. They should also be trained to accurately record responses.

The amount of training required will depend on who conducts the interviews. Experienced or professional interviewers may be familiar with features of complex questionnaires such as coding responses and skip patterns, and need less training to effectively administer the questionnaire. But peer interviewers drawn from the target population may be perceived as less judgemental. Although peer interviewers from the target population may need more training, they may also be perceived as less judgemental.

Training should include substantial role-playing so that interviewers become familiar and comfortable with the questions and skip patterns. Supervisors should also participate in this role play to become familiar with the interview process and to learn how to maintain quality control during survey implementation. More information on training can be found in Chapter A-16.

A-5.1.8 Pilot test the survey

After interviewers have been trained, the entire survey process should be pilot tested before actual data collection begins. This allows unforeseen challenges to be identified and corrected before the survey starts. The interview component should assess the feasibility of conducting the interview, the duration of the interview, and whether completed questionnaires and other survey instruments or forms are stored and transported as planned.

A-5.2 Summary

Data-collection instrument pretesting, questionnaire adaptation, interviewer training and survey procedure pretesting are all essential components of survey preparation and should be included in BBS protocols. Adequate advance time for these components needs to be incorporated into the survey preparation timeline.

To help ensure the quality of data collection and improve the survey instrument, investigators should:

- adapt data-collection instruments to the local context;
- develop and test software as warranted;
- ascertain the reading level of the target population to determine what interview mode should be used for self-administered questionnaires (CASI, ACASI or other);
- determine whether responses vary depending on interview mode – oral responses given during a face-to-face interview may differ from responses entered privately in a self-administered interview, and questions that are read (as in CASI) may be understood and answered differently than those that are heard (as in ACASI);
- take the time to translate the questions and then to back translate them into the original language, thus ensuring that complex concepts are interpretable in a commonly understood manner;
- check questionnaires – electronic or paper based – for logic and validity of response choices and skip patterns;
- introduce artificial errors and assess quality control by measuring the proportion of those errors identified during pretesting;
- pretest (electronic) questionnaires, and make improvements and corrections as needed;
- verify that the questions are clear to survey participants, understood as intended, and answerable; and
- examine the test data produced with these instruments.

A-5.3 References

1 Dillman DA. Mail and internet surveys: The tailored design method. Wiley New York. 2000.

2 Baral S, Logie CH, Grosso A, Wirtz AL, Beyrer C. Modified social ecological model: a tool to guide the assessment of the risks and risk contexts of HIV epidemics. BMC Public Health. 2013;13:482.

3 Hayes R, Kapiga S, Padian N, McCormack S, Wasserheit J. HIV prevention research: taking stock and the way forward. AIDS. 2010;24(Suppl 4):S81–S92.

4 Shannon K, Goldenberg SM, Deering KN, Strathdee SA. HIV infection among female sex workers in concentrated and high prevalence epidemics: why a structural determinants framework is needed. Curr Opin HIV AIDS. 2014;9(2):174–182.

5 Chersich MF, Luchters S, Ntaganira I, Gerbase A, Lo YR, Scorgie F et al. Priority interventions to reduce HIV transmission in sex work settings in sub-Saharan Africa and delivery of these services. J Int AIDS Soc. 2013;16(1):17980.

6 Mayer KH, Wheeler DP, Bekker L-G, Grinsztejn B, Remien RH, Sandfort TGM et al. Overcoming biological, behavioral, and structural vulnerabilities: new directions in research to decrease HIV transmission in men who have sex with men. J Acquir Immune Defic Syndr. 2013;63(Suppl 2):S161–S167.

7 Wheeler T, Kiran U, Dallabetta G, Jayaram M, Chandrasekaran P, Tangri A et al. Learning about scale, measurement and community mobilisation: reflections on the implementation of the Avahan HIV/AIDS initiative in India. J Epidemiol Community Health. 2012;66 Suppl 2:ii16–25 (http://www.ncbi.nlm.nih.gov/pubmed/22945907, accessed 4 August 2016).

8 Drobnik A, Pinchoff J, Bushnell G, Ly S, Yuan J, Varma JK et al. Matching HIV, tuberculosis, viral hepatitis, and sexually transmitted diseases surveillance data, 2000–2010: identification of infectious disease syndemics in New York City. J Public Health Manag Pract. 2014;20(5):506–512 (http://www.ncbi.nlm.nih.gov/pubmed/24335712, accessed 4 August 2016).

9 Russell BS, Eaton LA, Petersen-Williams P. Intersecting epidemics among pregnant women: alcohol use, interpersonal violence, and HIV infection in South Africa. Curr HIV/AIDS Rep. 2013;10(1):103–110 (http://www.ncbi.nlm.nih.gov/pubmed/23233038, accessed 4 August 2016).

10 Santos GM, Do T, Beck J, Makofane K, Arreola S, Pyun T et al. Syndemic conditions associated with increased HIV risk in a global sample of men who have sex with men. Sex Transm Infect. 2014;90(3):250–253 (http://www.ncbi.nlm.nih.gov/pubmed/24431183, accessed 4 August 2016).

11 Sudman S, Bradburn NM. Asking questions: A practical guide to questionnaire design. San Francisco, California, USA, Jossey-Bass. 1989.

12 Sudman S. Estimating response to follow-ups in mail surveys. Public Opin Q. 1983;46(4):582–584 (https://www.ncbi.nlm.nih.gov/pubmed/10278173, accessed 1 November 2016).

13 Mooney G. In: Lavrakas PJ (ed), Encyclopedia of survey research methods, Thousand Oaks, California, SAGE Publications. 2008.

14 Willis GB. Cognitive interviewing: a tool for improving questionnaire design. Thousand Oaks, California, SAGE. 2005.

15 Campanelli P. Testing survey questions: New directions in cognitive interviewing. Bulletin of Sociological Methodology. 1997;55(1):5–17 (http://bms.sagepub.com/content/55/1/5, accessed 30 August 2016).

16 Collins D. Pretesting survey instruments: an overview of cognitive methods. Qual Life Res. 2003;12(3):229–238.

6. Biomarker considerations

This chapter focuses on selecting biomarkers for measurement in biobehavioural surveys (BBS) and the laboratory considerations most relevant to them. It discusses everything from selecting biomarkers that are appropriate to the target population and local setting, to the practicalities of dealing with specimens.

Key terms

Algorithm:	Step-by-step procedure.
Biological data or "biodata":	the data derived from testing biological specimens.
Biological specimen:	Biological material, such as blood, urine or saliva, collected from a person.
Biomarker:	A biological characteristic measured through laboratory or other tests, generally using a biological specimen such as blood for testing; examples include HIV antibody or viral load.
Biosafety:	Containment principles, technologies and practices implemented to prevent unintentional exposure to pathogens and toxins, or their accidental release.
Biosecurity:	Protection, control and accountability practices for biological materials within laboratories that prevent their unauthorized access, loss, theft, misuse, diversion or intentional release.
Dried blood spot (DBS):	Small volume of blood, collected from a participant via a finger prick or other means, spotted onto a filter paper card, dried and transferred to a laboratory for analysis.
Plasma:	Liquid portion of blood in which red and white blood cells, platelets, nutrients, waste products, antibodies, clotting factors, hormones and proteins are suspended. When anticoagulated whole blood is centrifuged (or allowed to settle), red cells go to the bottom, and white cells and platelets to the middle, leaving the yellowish plasma at the top.
Quality assessment:	A set of activities conducted as part of quality assurance to evaluate a laboratory or testing site.

A-6.1 Biomarker selection

Survey investigators should choose biomarkers based on the biological information needed to characterize the local epidemic and risk factors for HIV (see Table A-6.3). Linkage of biomarkers with participant characteristics and their reported behaviours allows investigators to explore associations between them (e.g. HIV status by age group), and provides useful information for policy and service planning. Biomarker data collected in repeated surveys may be used to monitor trends in biomarker prevalence, which may then be explained using linkage with behavioural data. Because biological testing technology is constantly evolving, survey investigators should consult with laboratory advisors, national authorities, the Centers for Disease Control and Prevention (CDC), WHO and test kit manufacturers for the most up-to-date information regarding biomarker assays; specimen, instrumentation and commodity requirements; and SOPs. Collaboration with laboratory experts and the national reference or public health laboratory is strongly recommended, and should be established during the survey design phase.

Biomarker data may be qualitative (e.g. a categorical outcome, such as positive or negative) or quantitative (e.g. a measure on a continuous scale, such as 10 000 viral copies/mL). Going beyond measuring clinical and molecular biomarkers for estimating HIV prevalence, some markers (e.g. CD4+ T-cell count and viral load) are particularly useful for staging the HIV infection, measuring the potential for HIV transmission, and assessing viral suppression or treatment failure to measure antiretroviral therapy (ART) programme impact.

Biological specimens and biomarkers

The terms "biological specimens" and "biomarkers" are sometimes used interchangeably, but in fact have different meanings. In this document, "biological specimen" refers to the actual biological material collected from a survey participant, such as blood, urine or saliva. A "biomarker", often derived from a biological specimen through testing, is a measurable factor associated with health or a particular medical condition. Examples include HIV antibodies or viral load as markers of HIV infection; C-reactive protein, a marker of acute inflammation; CD4+ T-lymphocyte count, an indicator of immune function; and hepatitis B virus (HBV) surface antigen or antibodies to HBV core and surface antigen (see Appendix I-12). Biomarkers may also include genetic test results, such as HIV drug resistance or HIV subtype *(1)*. Some biomarkers do not require biological specimens but are derived from other measures (e.g. blood pressure to assess hypertension or the ratio of height to weight to identify malnutrition in children under five years of age).

When selecting biomarkers for inclusion in a BBS, investigators should assess the need for biological information about a population, taking into account previously collected data, and the information needs of different disease control programmes. For example, sexually transmitted infection (STI) programmes may be interested in measuring the prevalence of syphilis in men who have sex with men (MSM), and national hepatitis programmes may be interested in the rate of hepatitis C virus (HCV) among people who inject drugs (PWID).

A-6.1.1 Target population and local context

The appropriate biomarkers to collect during a BBS will depend on the target population and the local context. For example, STI biomarkers could be a higher priority for surveys among MSM, sex workers (SW) and transgender persons (TG) than among PWID. Testing for HCV could be a higher priority for a survey among PWID than among SW. However, HCV testing may be warranted where injecting drug use is common among SW. Also important to consider is the cost of tests and the local capacity for conducting them, and the feasibility and legality of exporting samples if testing can only be done outside of the country. When selecting biomarkers and laboratory methods, investigators should consider, but not be limited by, comparability with previous surveys. For example, as testing technology improves, investigators should consider using improved methods that may offer greater sensitivity and specificity rather than using only the methods employed in previous surveys. When interpreting results, however, investigators need to consider the impact that switching methods may have on analysis of differences or trends (*2*). Furthermore, investigators may decide to discontinue collection of biomarker data that are no longer needed or used.

Key terms

Quality assurance: A range of activities that enable laboratories to achieve and maintain high levels of accuracy and proficiency. Quality assurance:

- establishes standard operating procedures (SOPs) for each step of the laboratory testing process, from specimen handling to instrument performance validation;
- defines administrative requirements, such as mandatory record-keeping, data evaluation and internal audits to monitor adherence to SOPs;
- specifies corrective actions when problems are identified, documentation, and the persons responsible for carrying out corrective actions; and
- sustains high-quality employee performance.

Serum: Fluid that rises to the top when coagulated (clotted) blood has been centrifuged or allowed to separate. Clearer than plasma because it contains fewer proteins such as clotting factors, which are held in the clot at the bottom of the tube.

Venipuncture: Puncture of a vein with a needle to collect blood (also referred to as phlebotomy) or to administer intravenous therapy.

Viral load: Concentration of HIV in the blood expressed as the number of viral copies per millilitre. In clinical care, viral load measurement is used to monitor the effectiveness of antiretroviral treatment in suppressing viral load.

Whole blood: Blood that includes both the cellular elements (red and white blood cells), platelets and plasma.

A-6.1.2 Data needed for action

Investigators need to consider how the biological data will contribute to local knowledge, and whether the country or programme has the ability to act on the data obtained. Examples of how biological and behavioural data may be combined to assess need for public health action are shown in Table A-6.1.

Table A-6.1 Surveys that collect both biological and behavioural data may indicate levels of risk, and highlight the need for public health action

Biomarker and behavioural data	May indicate
HIV or STI biomarkers and data on risk behaviour	Risk for HIV acquisition or transmission
HIV viral load and data on risk behaviour	HIV transmission potential
CD4 count and HIV testing history	Delay in testing; eligibility for treatment
HIV status and data on receipt of prior HIV test results	Awareness of HIV serostatus
HBV infection status and data on injection behaviour	Risk for HBV acquisition or transmission, and potential need for immunization

A country may not have the resources to immediately act on survey findings that indicate high levels of risk behaviour or disease burden (e.g. by providing hepatitis treatment or vaccination). Nevertheless, the data may provide the foundation to advocate for implementing or strengthening prevention, care and treatment efforts.

A-6.2 Resource availability

Investigators must consider available resources, such as budget, laboratory infrastructure and human capacity when selecting which biomarkers to include in a BBS. If specimens can be stored in a repository, investigators may be able to perform additional testing or facilitate future shipment to an external laboratory when funding becomes available. Collecting, handling, transporting and storing specimens may have significant costs in addition to the costs of biomarker test kits, reagents and equipment. Investigators also must consider the human resources required for specimen collection, processing and testing, as well as for providing counselling, and referrals to care and treatment.

During the survey planning phase, survey investigators should collaborate with laboratory experts to determine local capacity for testing and storing specimens, where tests will be performed, and what testing platform and assays will be used.

A-6.3 Ethical considerations

As part of the informed consent process, participants must be told what specimens are being collected and how; which tests will be performed; the risks and benefits of testing; if, when and how test results will be returned; and which treatment or referral services will be provided. Also, when applicable, survey investigators must obtain informed consent to store specimens for a specified or indefinite period, and conduct unspecified future tests on the specimens. The survey protocol should specify the type and quantity of specimens to be collected (e.g. 10 mL of whole blood or 15 dried blood spots), amount to be stored (e.g. volume or number of spots), and the length of time specimens will be stored. It is generally considered unethical to obtain consent for future unspecified testing if investigators know at the time of consent what future tests they will conduct.

Protocols that seek to identify bacterial STIs must include a plan to provide treatment or referral to a health facility willing to provide treatment to survey participants. Protocols should also describe how investigators will measure successful referral of participants to HIV or STI care and treatment (e.g. participants may be given vouchers to present at referral sites). Vouchers can be collected and counted to measure the proportion of referred participants who presented to referral sites. In practice, it can be difficult to ensure successful referrals, and even harder to document them when they do occur. Documenting successful referrals may be particularly difficult because there may be multiple clinics from which participants may choose to receive care, including clinics not included on the referral list; also, some participants may leave the area or simply be lost to follow-up.

A-6.4 Providing test results to participants

Survey investigators should give participants all test results that could affect their clinical care. Results that are of no health consequence to the participant (e.g. illicit drug metabolites detected in urine, or prostate-specific antigen in vaginal swabs), or that have utility at the population level but limited accuracy at the individual level (e.g. recent HIV infection), may not need to be returned to the participant. The ethical obligation to provide test results and connect survey participants to care is especially important for key populations who

may have less access to health services than members of the general population. Where health services for key populations do not exist, investigators should identify and train willing providers to work with these populations.

Surveys can return results while maintaining the anonymity of participants by returning results either on the same day the specimen is collected or through a system whereby participants can anonymously verify their identities (e.g. with a unique code or through a coded fingerprint scan) and collect the result at a later date. Surveys that return results days or weeks after sample collection often find that some participants cannot be reached or do not return for their results, meaning that those participants miss an opportunity to learn their HIV or STI status. Rapid testing is thus ideal for ensuring that participants obtain results. Additionally, it may be logistically challenging to return results of some assays that are performed off-site or at a much later date (e.g. testing for HIV drug resistance). One option for making such results available to participants includes giving them a phone number that they can call after a specified period to obtain the results. This may be feasible even after the survey has closed.

A-6.5 On-site or off-site testing

Surveys that use rapid tests and other point-of-care technology to test for biomarkers on-site (or near the site of specimen collection) should provide participants with same-day results (e.g. for HIV, CD4+ T-cell count and syphilis). Investigators are strongly encouraged to use the national standard for diagnostic testing rather than a testing algorithm recommended only for surveillance purposes. If diagnostic testing is performed, the counsellor may be able to provide the participant with a final result based on rapid testing. Some national standards, however, may require that rapid test results be confirmed with another test that cannot be performed on-site. In this case, counsellors will need to ask the participant to return to obtain the final result or to visit another facility for confirmatory testing. For some biomarkers, including HIV, it is important to confirm positive or reactive test results by including a specific test (e.g. western blot or similar) to ensure that accurate results are returned to participants.

For specimens that are tested off-site (at a central or reference laboratory) a unique code or identifier can be used to return results to participants together with post-test counselling at a later date, either at the survey site or at nearby health facilities. A disadvantage of this approach is that participants may forget or decide not to return for their results, or may be unavailable during a repeat visit in the case of time-location sampling. Investigators could potentially use technologies such as mobile phones to return test results or send reminders, but their use implies the collection of personal identifiers that may complicate data security procedures. A material transfer agreement (Appendix I-11) should be used when specimens are sent off-site for testing.

A-6.6 Treatment and referral

If survey investigators test for HIV, they assume responsibility for returning results to participants and referring them to HIV services. Successful linkage of PLHIV to treatment is extremely important ethically and to combat the epidemic. Referral IDs linked to survey IDs may be used to track successful linkage to health or other services. The use of referral IDs can prevent service providers from connecting patient names to survey IDs. To maintain participant anonymity, delete referral IDs from the dataset prior to releasing a public dataset. Ensure that the names of people who were successfully referred are not communicated to survey staff.

If STI biomarkers are assessed, investigators should provide STI treatment or referral to treatment, as recommended by local health authorities, based on test results. Providing on-site treatment for bacterial STIs is preferable to referral because some participants may not follow through with the referral. For some STIs, such as herpes simplex virus-2 (HSV-2) infection, treatment may not be feasible or easily available. Special consideration should be paid to treatment of participants who are pregnant or have allergies, mode of drug administration (oral, intramuscular or intravenous), managing allergic reactions and length of treatment (single or multiple doses). If national treatment guidelines do not exist, investigators should consult the most recent WHO or other international treatment guidelines.

A-6.7 Testing for HIV biomarkers

HIV biomarkers include viral components (e.g. proteins, commonly referred to as HIV antigens) and nucleic acids (e.g. RNA and DNA) that indicate presence of HIV, and antibodies produced as part of the immune response to HIV infection. Antibody- and antigen-based assays are the most commonly used HIV diagnostic tests. HIV genomic sequencing is used for HIV drug resistance or molecular epidemiology. CD4+ T-cell assay results indicate level of immunosuppression. Viral load is an indicator of virus transmission potential and the effectiveness of antiretroviral treatment. Providing additional test results, such as CD4+ T-cell count and viral load, to HIV-infected survey participants together with a referral to the nearest health facility offering ART may facilitate linkage of participants to treatment.

Table A-6.2 lists commonly measured HIV-related biomarkers and their utility.

Table A-6.2 HIV-related biomarkers and tests, and their usefulness in BBS

Biomarker	Test	Usefulness in BBS
HIV antibody		Presence indicates HIV exposure and infection; HIV seroprevalence is the proportion of the population with HIV antibody. Can be used to validate self-reported HIV status and measure the proportion of sample population undiagnosed or unaware of their status. An exception is infants who may have maternal antibodies to HIV but who are not themselves infected with HIV.
	RT	Rapid testing facilitates provision of same-day test results. Some RTs also detect HIV antigens, while others can distinguish between recent and long-term infection.
	EIA	Laboratory-based method that is more efficient than RTs for testing a large number of batched specimens, but same-day results may not be feasible.
	WB	Confirms EIA results (if required by national guidelines). Other highly specific rapid confirmatory tests are now available that may be used to confirm positive results, HIV-1 and HIV-2 dually reactive specimens, or serve as a tiebreaker for discordant results, and can discriminate between HIV-1 and HIV-2 infections.
	Tests for recent HIV infection	Distinguish recent from long-term HIV-1 infection to estimate population-level incidence; tests are performed on HIV-1-seropositive specimens and the testing algorithm should include viral load testing to improve the predictive value of a test for recent infection. These tests are used to estimate HIV incidence at the population level, and usually are not recommended for individual diagnosis. Most BBS cannot achieve sample sizes large enough to estimate assay-based HIV incidence with adequate precision.
HIV antigen (e.g. p24 antigen)		Presence indicates HIV infection earlier than antibody detection.
	WB	Rarely used in surveys.
	RT	Combination assays that independently distinguish results for HIV-1 p24 antigen and HIV antibodies in a single test are now commercially available.
CD4+ T-lymphocytes	CD4+ T-cell count	Indicates level of immunosuppression and ART eligibility; approximates duration of infection in treatment-naive individuals. CD4 testing requires a whole blood specimen, which should be processed within 24 hours of collection.
HIV RNA	VL test	Measures amount of virus in the blood. Indicator of treatment effectiveness and transmission potential; can be used to determine population (summary) VL and proportion of HIV-infected persons with suppressed (e.g. <1000 copies/mL) VL. Definition of viral suppression may vary based on factors such as clinical guidelines and assay used. Common definitions include: • population VL: among those HIV+ (diagnosed or not) • diagnosed VL: among those diagnosed with HIV • treatment VL: among those on ART • indicator for final step of HIV cascade: percentage on ART who are virally suppressed.
HIV drug resistance (HIVDR) mutations	Genomic sequencing	Indicates level of resistance to different ARVs. HIVDR may be reported and used at the individual level to help guide patient treatment, or at the population level to estimate levels of resistance to different classes of ARVs.
HIV subtype	Genotyping	Identifies clusters of HIV infection in a population; useful in describing subepidemics (e.g. when HIV subtypes among key populations differ from those in the general population).
ARV metabolites	HPLC combined with mass spectrophotometry	Detects presence of ARV metabolites. Marker of ART or use of PrEP. May validate clinical or self-reported information about ARV use (e.g. ART adherence, or whether or not participant is currently on ART). Percentage on ART is an HIV cascade indicator.

ART, antiretroviral therapy; ARV, antiretroviral; EIA, enzyme immunoassay; HIVDR, HIV drug resistance; HPLC, high-performance liquid chromatography; PrEP, pre-exposure prophylaxis; RT, rapid test; VL, viral load; WB, western blot

A-6.8 Testing for non-HIV biomarkers

BBS often include other biomarkers, particularly for STIs. Many STIs are transmitted much more easily than HIV, and such infections are often easily treated and less likely to become chronic. Thus, STI prevalence is often a more sensitive indicator of recent sexual risk behaviour than HIV prevalence. Some STI measurements in surveys can serve as:

- an early warning system for potential HIV acquisition or transmission and emergence of HIV in new groups or geographical areas;
- a risk assessment tool for HIV prevention programmes; and
- a means to detect independent STI epidemics (e.g. syphilis or gonorrhoea) that need public health attention.

Hepatitis B (caused by HBV) and hepatitis C (caused by HCV) are bloodborne infections that are easily transmitted through contact with blood, including by sharing injecting equipment, and occasionally through sex. These infections, as well as other viral STIs such as HSV-2, have a tendency to become chronic infections, and treatment may be difficult or expensive to obtain, or the infection may not be curable. Measurement of these biomarkers is less useful for correlation with sexual risk behaviours, although prevalence of HBV and HCV may be useful indicators for monitoring the impact of prevention programmes among drug users. Measures of HSV-2 may be useful for estimating the lifetime exposure to unprotected sex, especially among young people.

Table A-6.3 lists commonly measured non-HIV biomarkers and their utility.

Table A-6.3 Non-HIV biomarkers and tests, and their usefulness in BBS

Non-HIV biomarker	Test	Usefulness
***Chlamydia trachomatis* (CT)**		Causative agent of chlamydia. All listed tests detect active infection with CT. Useful for validating self-reported sexual behaviour; monitoring trends in sexual risk behaviour; and measuring prevalence of CT in the survey population.
	Nucleic acid amplification test (NAAT)	NAAT is highly sensitive and specific and requires closed systems and trained laboratory staff. Involves noninvasive specimen collection (urine or swabs – vaginal or rectal). Current standard of testing in high-income countries but increasingly also used in resource-constrained settings. Many NAATs include both CT and NG probes allowing testing for both organisms with a single specimen. For surveillance purposes (e.g. BBS) NAAT should be considered over other test methods.
	Direct fluorescent antibody (DFA)	DFA was previously used for clinical diagnosis, but is currently not recommended for routine testing of genital tract specimens. Procedure requires experienced microscopist and is labour intensive and time consuming. Not useful for BBS.
	Enzyme immunoassay (EIA)	EIA is used in clinical diagnosis. Detects chlamydial antigen. Requires experienced technician and collection of adequate specimen. Also requires confirmatory assay (with a blocking antibody or direct fluorescent antibody test) to eliminate false positive result. None of the EIAs are as sensitive or specific as the NAATs, and costs are typically high. Impractical for BBS.
	Cell culture	Although culture was formerly the reference standard against which all other tests have been compared, few laboratories have capacity for culture of this delicate organism. Culture methods for C. trachomatis are insensitive, difficult to standardize, technically demanding, and expensive. Maintaining viability of organisms during transport and storage in diverse settings is a serious challenge. Impractical for BBS.
	Serology	Current serological tests have limited if any value for screening due to cross-reactive antibodies.
***Neisseria gonorrhoeae* (NG)**		Causative agent of gonorrhoea. All tests listed detect active infection with NG, indicating recent unprotected sexual intercourse. Useful for validating self-reported sexual behaviour; monitoring trends in sexual risk behaviour; and measuring prevalence of NG in the survey population.

Non-HIV biomarker	Test	Usefulness
	NAAT	NAAT is highly sensitive and specific and requires closed systems and trained laboratory staff. Involves noninvasive specimen collection (urine or swabs – vaginal, rectal or oropharyngeal). Current standard of testing in high-income countries, increasingly being used in resource-constrained settings. Many NAATs include both CT and NG probes allowing testing for both organisms with a single specimen. For surveillance purposes (e.g. BBS) NAAT should be considered over other test methods.
	Gram stain or gentian violet stain (GS or GV)	GS or GV is useful for diagnosis of symptomatic disease in men with urethritis, but not commonly used in BBS. Useful for distinguishing NG from CT in symptomatic men through detection of intracellular diplococci. Less useful in women, and not useful in asymptomatic men or women. Requires microscope and experienced technician.
	Culture	Culture had been the reference standard against which all other tests have been compared; however, NG is an obligate anaerobe and maintaining the viability of organisms during transport and storage in the diverse settings is a serious challenge. Cultures require enriched media. Because of the difficulty in transporting NG, culture is not very sensitive (i.e. most of the organisms die during transport). Culture is currently needed for studying antibiotic resistance.
	DFA	Detects NG antigen in genital secretions. Same caveats as for CT. DFA testing is difficult, expensive and not very sensitive.
	EIA	See above caveats for CT. EIA is less sensitive than NAAT and not recommended.
	Serology	There is no serological assay available for NG.
Treponema Pallidum (T. pallidum)		Causative agent of syphilis
	Serological tests – nontreponemal and treponemal antibodies	Serological tests may be used to detect current or past syphilis. Nontreponemal antibodies may indicate active (current) syphilis. Treponemal antibodies may indicate either active or past (resolved) syphilis. Specimens reactive for nontreponemal antibodies are usually confirmed by testing for treponemal antibodies. Useful for estimating prevalence and for diagnosis and treatment of syphilis among key populations. Infection with nonsexually transmitted treponematoses (e.g. yaws and pinta) will result in a positive serological test for T. pallidum and should be considered where such infections are endemic. Currently, there are no routine laboratory methods to distinguish endemic treponematoses from each other or from syphilis.
	Nontreponemal tests – indirect method that detect biomarkers released during cellular damage caused by *T. pallidum*. These include: • VDRL (venereal disease research laboratory) slide test • Unheated serum reagin (USR) test • RPR (rapid plasma reagin) card test • TRUST (toluidine red unheated serum test) • rapid immunoassays	Can be useful for screening; rapid, simple and inexpensive, but require subjective interpretation by an experienced laboratory technician. Nontreponemal tests can be used as a marker of sexual activity. Reactive results may be confirmed by a treponemal test (see below). Because one of the objectives of BBS may be to detect active infections, reactive nontreponemal tests should be confirmed with a treponemal test. Likewise, reactive treponemal test results should be confirmed with a nontreponemal test (e.g. RPR or VDRL) to confirm current syphilis infection. VDRL and USR must be read under a microscope. RPR and TRUST do not require a microscope, but do require trained personnel using standardized equipment and specialized reagents.

Non-HIV biomarker	Test	Usefulness
	Treponemal tests – detect *T. pallidum* antibodies that occur as a direct result of infection • TPPA (*T. pallidum* particle agglutination) • FTA-Abs (fluorescent treponemal antibody absorption) • TPHA (*T. pallidum* particle haemagglutination assay) • Rapid immunoassays • Combination rapid assays (treponemal and nontreponemal)	A reactive treponemal test indicates current or past infection and may not differentiate between the two. Treponemal test results may remain reactive for life, even with treatment. Because a reactive test result only indicates exposure to TP at some time in a person's life, it needs to be confirmed with a nontreponemal test. Treponemal tests do not differentiate venereal syphilis from endemic syphilis (yaws and pinta). Traditionally treponemal tests were used mainly as confirmatory tests to verify reactive screening test results. However, a rapid treponemal test or EIA can be used for screening in certain situations (e.g. population prevalence <1% or screening in antenatal women). Individuals with reactive test results would then be treated presumptively or have a follow-up RPR or VDRL to identify active infection. Rapid immunoassays are available that simultaneously test for treponemal and nontreponemal antibodies, simplifying the detection of both current and past infection.
	Dark-field microscopy	Useful for diagnosis when lesions are present, but impractical for BBS. Requires examination by trained technicians immediately (within 20 minutes) after specimen is collected onto a slide, and a microscope with dark-field illumination. Dark-field examination of lesion exudates for motile spirochetes may be helpful, but if a patient has self-medicated or if there are only a few organisms present, spirochetes may not be seen.
Trichomonas vaginalis (T. vaginalis)		Causative agent of trichomoniasis, the most prevalent nonviral STI; trichomoniasis prevalence.
	Direct microscopic examination (DME) – wet-mount preparation of vaginal discharge to assess presence of motile protozoa	DME is inexpensive and easy; however, sensitivity is highly variable (38% to 82%) and is dependent upon inoculum size and skill of examiner. Requires cervical smears, and thus often impractical for BBS unless an examination is conducted.
	Nucleic acid detection (Polymerase chain reaction-PCR)	Several PCR-based diagnostic assays using vaginal and urine specimens exist; these vary in their sensitivity and specificity. Can be useful in BBS using urine (men) or vaginal swab (women) samples.
	Broth culture	Sensitivity varies by culture technique. Less expensive and more convenient than cell cultures, but insensitive compared with PCR. Could be used for BBS, but may be costly.
Haemophilus ducreyi		Causative agent of chancroid; facilitates HIV transmission and acquisition; extremely delicate organism and thus difficult to culture. Clinically, chancroid can appear similar to syphilis in causing genital ulcer syndrome; however, chancroid is typically painful and associated with more lymphadenopathy than syphilis.
	Nucleic acid detection with or without amplification	See description of PCR above.
	Serology	Serological tests have limited sensitivity in individuals, but may be useful in population studies.
	DFA (direct fluorescent antibody)	See description of and caveats about DFA above.
	Culture	Culture had been the reference standard against which all other tests have been compared; however, with development of PCR it has been realized that culture is only about 75% sensitive at best. H. ducreyi is difficult to transport and culture. Cultures require highly enriched media. Not practical for BBS.

Non-HIV biomarker	Test	Usefulness
Herpes simplex virus-2 (HSV-2)		HSV-2 is a causative agent of genital herpes and a cofactor of HIV infection; indicator of population risk for ever having unprotected sex. Estimation of HSV-2 prevalence may be most useful for surveys in young age groups.
	EIA (for HSV antibody) and other serological tests	EIAs can aid in the diagnosis of HSV: a positive result suggests recent infection, but cannot distinguish between primary or reactivated infection. A negative result does not rule out infection. Costly and difficult; not useful in screening. Several type-specific serological assays exist, although some cross-reactivity may occur between HSV-1 and HSV-2. Serology is preferred for screening populations and tests perform reasonably well on a population level compared with individual diagnosis.
	DFA (direct fluorescent antibody)	DFA testing requires scraping cells from the ulcer base in symptomatic patients. Can be useful in distinguishing HSV-1 from HSV-2, but not practical for surveys.
	PCR	PCR is sensitive and specific in symptomatic patients, and can be used to detect asymptomatic viral shedding in infected persons. Not useful for screening.
	Western blot	Rarely used.
	Viral culture	Costly and difficult. Specimen collection must be done during acute phase of infection when skin lesions are present. Not practical for surveys.
Hepatitis B virus (HBV)		HBV prevalence – indication of acute or past infection, chronic carriage, or immunization depending on tests used (see Appendix I-12); prevalence of HIV-HBV coinfection. Assess impact of prevention, care, control and treatment programmes.
	EIA	HBsAg, Anti-HBc, IgM anti-HBc, Anti-HBs
	Nucleic acid testing	HBV DNA
Hepatitis C virus (HCV)	HCV RNA	HCV prevalence; common among PWID; prevalence of HIV-HCV coinfection. Assess impact of prevention, care, control and treatment programmes. May be used as indicator of injecting drug use, but need to understand current and historical modes of HCV transmission in the local setting. With the availability of HCV treatments, monitoring the HCV testing, care and treatment cascade is becoming more important.
	EIA or rapid diagnostic tests (RDTs)	Detect anti-HCV antibody (cannot distinguish new or reinfection, chronic infection or past resolved infection).
	HCV RNA Sputum-smear microscopy	Biomarker for viremia; indicates current or chronic HCV infection; distinguishes those with current infection from those whose infection has resolved
Mycobacterium tuberculosis		Causative agent of tuberculosis (TB); prevalence of TB or TB disease among those with HIV; TB disease among those with HIV indicates severe immunosuppression and is an AIDS-defining condition. Most BBS cannot achieve sample sizes large enough to estimate TB incidence or prevalence in a population, but testing may be appropriate in some instances, such as in BBS conducted among prisoners.

Non-HIV biomarker	Test	Usefulness
	• Sputum smear microscopy: conventional light microscopy used to examine direct smears stained with Ziehl–Neelsen, with or without specific sputum-processing methods; or • Fluorescence microscopy; or • Rapid molecular test with high accuracy, such as the Xpert MTB/RIF test (Cepheid, Sunnyvale, CA); or • Any rapid test recommended by WHO	• Sputum smear microscopy (including fluorescence microscopy) is highly specific in the diagnosis of pulmonary TB among persons in areas with a very high prevalence of TB and can identify those who are most infectious. Sensitivity of detection is low, however, for sputum specimens with fewer than 10,000 bacteria/mL specimen and among those with extra-pulmonary TB and in persons with HIV-TB coinfection. • Molecular tests provide rapid results for diagnosis of TB. The Xpert MTB/RIF assay contributes to the rapid diagnosis of TB disease and rifampin drug resistance.
Prostate-specific antigen	Immunochromatographic rapid test on vaginal or rectal swabs	A marker of seminal fluid indicating recent unprotected sexual intercourse.
Y-chromosomal sequences (YCS)	Qualitative PCR assay	A marker for unprotected sex among females. The detection of YCS or lack thereof determines the presence or absence of male chromosomes in the vaginal swab sample. The assay has been shown to be 92–98% specific for YCS detection within 2–3 weeks of unprotected heterosexual intercourse (3).
Illicit drugs (e.g. amphetamines, cannabis, cocaine, opiates)		• Validate self-reported drug use; estimate proportion of population using drugs, by type. • Tests may use blood, urine or hair, with detection times (since last drug ingestion) varying by drug and specimen type. • Rapid tests exist that show reasonable accuracy and can test concurrently for multiple analytes

Anti-HBc, hepatitis B core antibody; Anti-HBs, hepatitis B surface antibody; HBsAg, hepatitis B surface antigen; IgM, immunoglobulin M; TP, T. pallidum

A-6.9 Selection of tests and testing algorithms

Biomarkers are among the most important survey data collected, because they typically relate to a survey's primary and secondary objectives. Table A-6.3 provides information about various STIs and their tests that can be used in selecting STIs for surveys.

Appendix I-10 compares specimen types, advantages, limitations, cost and complexity of the two most commonly used HIV testing technologies, enzyme immunoassays (EIAs) and rapid tests. If participants are to receive HIV test results on the same day they participate in the survey, a rapid HIV testing algorithm will probably be required. The first test performed should be highly sensitive (to detect all true positives), and should be followed by a second highly specific test to identify any true negative specimens falsely identified as positive (i.e. false positive) by the first test. Some national algorithms require a third or confirmatory test. Additional guidance on designing HIV test algorithms for different surveillance purposes and in different epidemiological contexts is provided in the Joint United Nations Programme on HIV/AIDS (UNAIDS)/WHO *Guidelines for using HIV testing technologies in surveillance: selection, evaluation, and implementation – 2009 update (4).* These guidelines also provide information on the qualification of HIV and other tests to be used for diagnosis (5).

A-6.10 Development of standard operating procedures

Detailed SOP for sampling, transport, testing and data collection should be prepared for each test and type of specimen collected, ideally in consultation with laboratory technicians with expertise in the tests. Laboratory staff may adapt SOP from the package inserts of commercial test kits. SOP should include details about how specimens will be collected, processed, stored and transported, if applicable. Additionally, protocols and SOP should outline how and to whom test results will be returned. SOP also should address specimen disposal and adherence to biosafety guidelines. A sample SOP is provided in Appendix I-30. Training of survey staff should include exercises to identify and address unforeseen challenges.

A-6.11 Specimen type, collection, labelling, processing and transport

Survey biomarkers are commonly based on blood specimens (e.g. HIV, syphilis and HSV). Depending on the test, biomarker and target population, other specimens collected may include urine, oral fluid, vaginal swabs, rectal swabs and oropharyngeal swabs. Less commonly collected specimens (e.g. urethral swabs or endocervical swabs) may require a clinician to collect them; such specimens may be useful in symptomatic populations or among key populations at high risk for STI. Biosafety and biosecurity precautions must be strictly adhered to during specimen collection, handling, testing, storage and waste disposal, even where surveys are conducted in informal settings such as truck stops, bars and nightclubs, or rented apartments.

A-6.11.1 Specimen type and collection

Appendix I-10 provides a list of specimen types for HIV testing and the advantages and disadvantages of each. Table A-6.4 provides a list of optimal specimen types for the detection of *Chlamydia trachomatis and Neisseria gonorrhoeae* by nucleic acid amplification test (NAAT) type. Comprehensive information on laboratory methods for detecting *Chlamydia trachomatis* and *Neisseria gonorrhoeae* and other STIs can be found in *Recommendations for the laboratory-based detection of Chlamydia trachomatis and Neisseria gonorrhoeae (6),* and *Strategies and laboratory methods for strengthening surveillance of sexually transmitted infections (7).*

Blood

Blood – in the form of serum, plasma or dried blood spot (DBS) – is the preferred specimen for HIV testing because it has a higher concentration of HIV antibodies than oral fluid. It also allows for additional testing; for example, for viral load, recent infection, HIV type and subtype, and HIV drug resistance, as well as syphilis and hepatitis.

Blood can be collected through finger stick or venipuncture. Generally, participants are resistant to having more than one blood draw or finger stick (pricking). In addition to the discomfort, double-pricking may also compromise confidentiality because it may indicate that an infected person requires a second test. A single finger stick often yields only 200 µL of blood or less, although some systems may facilitate collection of up to 500 µL. Collecting venous blood using evacuated (vacuum) blood tubes is therefore more productive. With this approach, relatively large volumes (5–10 mL or more) of blood can be collected, meaning that multiple tests can be performed.

Collecting extra blood is useful in case tests have to be repeated for quality control, or to store specimens for future unspecified testing. The use of a needle and syringe and subsequent transfer to tubes is not encouraged because of biosafety issues and the potential for compromising the quality of the specimen (e.g. hemolysis).

Evacuated tubes containing the anticoagulant ethylenediaminetetraacetic acid (EDTA) are especially useful because they yield both cellular and noncellular (plasma) fractions. Vacuum tubes without anticoagulant (i.e. untreated) should be used if serum is desired. Aliquots of plasma or serum should be prepared shortly after collection, then refrigerated, transported and frozen at the central laboratory; alternatively, they may be frozen on-site and transported on freezer packs to the central laboratory without thawing. Repeated freezing and thawing should be avoided because it may compromise specimen quality.

Blood collected for CD4+ T-cell counting needs to be stored at ambient temperature. Special “long-life” blood tubes are available that stabilize the CD4+ T-cells up to 7 days, allowing for less stringent time requirements for specimen transportation and testing.

DBS are easy to prepare and they provide a backup specimen in case plasma specimens are lost or spoiled. However, DBS are not the ideal specimen type for HIV serology using EIA. Some EIA kit manufacturers have protocols for DBS but, in general, protocols have to be adjusted (optimized) because of the high background optical density readings associated with serum eluted from DBS. Moreover, DBS may not be suitable for every biomarker (i.e. CD4+ T-cells).

Table A-6.4 Blood specimens – collection, fractions and use

Blood specimen type	Collection container	Blood fraction	HIV rapid test	HIV EIA	HIV recency	Western blot	Viral load	HIV DNA PCR	HIV drug resistance	CD4+	Syphilis	HSV	Viral hepatitis	Other EIA
Venous blood	Vacuum, EDTA	Anticoagulated whole blood	✓							✓				
		Plasma	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
		Dried blood spots*		✓	✓	✓	✓	✓	✓					
	Vacuum, EDTA, preservative	Anticoagulated whole blood								✓				
	Vacuum, no anticoagulant	Serum	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Capillary blood	Capillary or microcollection tube	Anticoagulated whole blood	✓				✓	✓		✓				
		Dried blood spots*		✓	✓	✓		✓	✓					✓

* Prepared from whole blood on Whatman 903 filter paper cards

EDTA, ethylenediaminetetraacetic acid; EIA, enzyme immunoassay; HSV, herpes simplex virus; PCR, polymerase chain reaction

Table A-6.4 outlines type of blood specimens, methods of collection, and use for measuring HIV biomarkers. Additional guidance on the selection of specimen types is provided in the UNAIDS/WHO *Guidelines for using HIV testing technologies in surveillance: selection, evaluation, and implementation – 2009 update (4)*. For more specific information on specimen transportation and storage guidelines, refer to the WHO *Guidelines for the safe transport of infectious substances and diagnostic specimens (8)*. A laboratory assessment should be performed before BBS implementation to ensure quality standards, followed by implementation of a quality assurance and monitoring plan to assure testing quality during survey implementation (see Section A-6.15).

Oral fluid

Oral fluid (saliva) testing is less invasive than venipuncture, and can be useful for survey situations where it may be difficult to collect blood specimens. However, there are limitations to this approach. For example, if more than one oral fluid test is used, specimen collection must occur at least 30 minutes apart; also, oral fluid can only be used with certain EIAs and rapid tests specifically designed for oral fluid specimens. Oral fluid rapid HIV testing is sometimes performed in challenging survey environments where participants with reactive tests are counselled to seek additional testing for diagnostic confirmation elsewhere. Investigators may, however, choose to use the results for data analysis, without waiting for results of confirmatory tests.

In general, oral fluid specimens may be stored at 4°C to 37°C up to 21 days. This temperature should be maintained during shipment. Specimens should be frozen (–20 °C or below) if stored for an extended time. Once thawed, specimens can be refrozen and thawed only once more.

Urine

Urine is the CDC-recommended specimen for use in NAATs to detect *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in men *(6)*. Because the bacteria in first-catch urine specimens are remarkably similar to those found in paired urethral swab specimens, urine can be used when swabs are undesirable, such as in population-based studies or where multiple sampling of participants is required *(9)*. For BBS testing for *Chlamydia trachomatis* and *Neisseria gonorrhoeae*, survey investigators may choose to collect urine from men and from women;

however, because the sensitivity of vaginal swabs for detecting *Chlamydia trachomatis* and *Neisseria gonorrhoeae* among women is greater than that of urine, swabs are considered the optimal specimen type for women *(6)*.

Vaginal and endocervical swabs

For NAAT detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in women, vaginal swabs are considered the optimal specimen because they have greater sensitivity than first-catch urine specimens, and are as sensitive and specific as cervical swab specimens *(6)*. The sensitivity and specificity of self-collected vaginal swabs are equal to or better than swabs collected by a clinician.[1] Some women may find self-collection more acceptable. Also, self-collection reduces staffing requirements for the survey. Formative assessment should include questions about willingness to provide specimens and which collection method is preferred. Vaginal swabs do not replace cervical exams for diagnosis of female urogenital infection; women may have cervicitis, urethritis, or urinary tract or vaginal infections due to causes other than *Chlamydia trachomatis* or *Neisseria gonorrhoeae.*

Rectal swabs

For MSM, rectal swabs should be collected for NAAT detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. Depending on the context, rectal swabs may be collected among female SW (FSW) as well. These can be collected by a clinician or self-collected. As with vaginal swabs, formative assessment should probe the target population's willingness to provide rectal swab specimens and their preferred method of collection.

Oropharyngeal swabs

NAATs are the recommended testing method for detection of *Neisseria gonorrhoeae* in oropharyngeal specimens. Nonharmful *Neisseria* species commonly found in the oropharynx might cause false positive reactions in some NAATs, and further testing, such as by polymerase chain reaction (PCR), might be required for accuracy. *Chlamydia trachomatis* is not considered to be a clinically important cause of pharyngitis.

Urethral swabs

Urethral swabs historically have been the specimen of choice for STI diagnosis in men. Urethral swab collection, however, causes considerable discomfort and has been identified as a disincentive for routine STI screening *(4)*. Furthermore, urethral swab collection requires a clinician or trained health worker. Urine is the CDC-recommended sample type for NAAT-based diagnostics for *Chlamydia trachomatis* and *Neisseria gonorrhoeae (6)*. Urethral swabs are still recommended for testing men for *Trichomonas vaginalis*.

Table A-6.5 shows the ideal specimen method for each type of STI screening.

Table A-6.5 Optimal specimen types for BBS STI screening using nucleic acid amplification tests

Organism	Optimal specimen	Alternative	Reduced sensitivity
Women			
Chlamydia trachomatis	Vaginal swab	Endocervical swab	First-catch urine
Neisseria gonorrhoeae	Vaginal swab	Endocervical swab	First-catch urine
Trichomonas vaginalis	Vaginal swab Endocervical swab First-catch urine		
Men			
Chlamydia trachomatis	First-catch urine	Urethral swab	
Neisseria gonorrhoeae	First-catch urine	Urethral swab	
Trichomonas vaginalis	Urethral swab		
SW, MSM, TG (specimens in addition to those above)			
Chlamydia trachomatis	Rectal swab		
Neisseria gonorrhoeae	Rectal swab Oropharyngeal swab		

[a] Adapted from ARUP Laboratories 2014 (*10*)

A-6.11.2 Specimen labelling

All specimens (e.g. aliquots, DBS cards and rapid testing devices) must be labelled with the participant's unique identification number (ID) or a laboratory ID that is linked with the participant ID. The label is required for linking biological and survey data, and for specimen storage, tracking and retrieval. Careful planning of labelling procedures is recommended: a single 7 mL venous blood specimen drawn into a plain or anticoagulant tube may result in one or two 1.5 mL aliquots of serum or plasma and a five-spot DBS card, each of which require a label.

Labels should be preprinted before the survey begins. The number of labels per participant depends on the number of specimens and forms used. Extra labels should be printed for unanticipated needs. Ideally, labels are barcoded so that laboratory and survey staff can register specimens with a barcode reader to minimize manual data entry and associated errors. Preprinted labels that withstand humidity and freezing (down to at least –20 °C, and perhaps as low as –80 °C depending on where specimens will be stored) are the best option. Otherwise, a permanent marker should be used to label specimens. At a minimum, specimens should be labelled with the participant ID and date of collection. If not indicated on the specimens themselves, a record should be kept of which specimens may be stored for future testing and which must be discarded.

Temporary storage on-site and processing

Specimens may be stored temporarily in refrigerators or cooler boxes with cold or freezer packs. Specimen storage requirements vary by type of specimen. For example, DBS specimens should not be refrigerated immediately after collection, but held at room temperature and dried thoroughly before packaging. DBS should be packaged in a zip-lock bag containing a humidity indicator card and desiccant packets and stored in a refrigerator (4–8 °C) for up to a month or in a freezer (–20 °C) if available for long-term storage. If serum or plasma specimens will undergo any molecular testing (e.g. genotyping), they should be processed and frozen within 24 hours.

Handling, storage and transport of specimens collected for NAAT detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* testing varies by test kit manufacturer. Urine specimen storage requirements will vary by NAAT type and type of container (primary cup or transport tube). For example, urine is one type of NAAT that may be held in a primary cup at room temperature for up to 24 hours for female urine specimens and up to 3 days for male specimens; the holding time for other NAAT types ranges from 24 hours to 30 days at 2 °C to 30 °C *(6)*. Thermometers should be used to monitor storage temperature daily. Thermometers that record the minimum and maximum temperature are useful.

A-6.11.3 Transport

Specimens sent to the laboratory should be packaged and adequately stored during transportation (e.g. cooler boxes with or without ice packs, or liquid nitrogen), according to requirements of the assays to be performed. They should be accompanied by transportation or shipping forms, and chain-of-custody forms (Appendices I-13 and I-14). To ensure proper handling of specimens upon arrival, dates of shipping and receiving should be scheduled taking into account the receiving laboratory's operating hours. When coordinating local transport, investigators should consider whether the receiving laboratory will be staffed to process specimens at the time of their arrival. Blood collected in tubes with anticoagulant needs to be separated by centrifugation and the plasma frozen within 24 hours of collection for any molecular testing. In addition to being trained on proper specimen handling procedures, staff transporting fresh specimens should be trained on relevant biosafety and biosecurity procedures.

Specimen registration at the laboratory

When specimens are received by the laboratory, they should be cross-checked against the shipping form, registered into the laboratory data-management system (or logbook), assigned a unique laboratory ID that is linked to the participant ID (if applicable), and stored appropriately until used for testing. The laboratory data-management system should ensure specimen tracking and linkage of laboratory and participant IDs. Any nonconformity should be documented; for example, too high or low temperature, hemorrhage, bacterial overgrowth, empty or insufficient specimen volume, unsealed bags or missing labels.

Laboratory support

BBS should be supported by a laboratory, ideally one accredited by an international accreditation agency. Staff at the laboratory should participate in all phases of the survey from design to results dissemination (e.g. final report and manuscript writing). Countries without accredited laboratories may have a national reference laboratory, a national public health laboratory, or a university or hospital laboratory that may support surveys. The survey laboratory should have in place robust quality management systems and the capacity to perform at least basic diagnostic testing (e.g. serology testing). The laboratory may receive, process, store and test specimens, either to provide primary results for the survey, or to conduct quality-control testing or future testing. It may also provide support for on-site laboratory procedures by providing training, supplies and supervision.

A-6.12 Laboratory data management

Survey investigators should ensure that a system (preferably electronic) is in place for managing data; for example, cataloguing specimens, recording results, storing raw laboratory data and recording quality-control results. Survey laboratory results (e.g. line-listed optical density values, viral RNA copies/mL and quality-control test results) can be collected and merged with other survey data so that output for all the assays are maintained in the final database rather than just the final positive or negative interpretations. This will facilitate data review and troubleshooting, should testing issues arise.

A-6.13 Laboratory procedures

Even when testing of specimens is not being done at the survey site, BBS usually require on-site procedures for specimen collection and processing at the very least. Quality-assurance measures need to be considered, including training, use of quality-control materials, periodic review of testing data, proficiency testing panels, and retesting of a subset of specimens at a designated reference laboratory. Traditionally, HIV and STI testing have been conducted within the confines of a standard laboratory; however, there is a clear trend towards conducting on-site rapid testing without the need for sophisticated equipment, and inclusion of point-of-care testing (e.g. for CD4+ T-cells counts and viral load) with increasingly small and portable computerized equipment. A private room – or other designated location in the case of time-location sampling (TLS) – should be reserved for on-site laboratory procedures.

On-site infrastructure may include refrigerators or cooler boxes (for temperature-sensitive reagents or specimen storage); sharps, biohazard and biowaste disposal containers; centrifuges; mobile phones; and tables and chairs. Appendix I-2 provides examples of various cost categories, including a list of laboratory commodities and equipment.

A-6.14 Specimen repository

Survey protocols and consent forms often include text to allow for long-term storage of specimens for future unspecified needs. Protocols should specify a minimum and maximum storage period, and define who will control the use of the specimens. Investigators should consult relevant ethics review committees about the future use of specimens. Storage space is another important consideration. Many laboratories have limited space; thus, storage space needs to be secured in advance.

To store leftover survey specimens for future unspecified testing, survey investigators must obtain informed consent from survey participants. Participants must also be informed as to whether test results from future testing will be returned to them. The survey protocol should specify the type (e.g. plasma or DBS) and amount (e.g. volume or number of spots) of specimens to be stored. If a specimen repository is important to the BBS objectives, then ideally, two sets of specimens should be prepared – one for immediate testing, the other for storage and future testing.

Long-term specimen repositories must be continuously managed. For example, DBS specimens must be inspected every 6 months and their desiccants replaced as necessary. Use of a centralized laboratory information system may help with tracking and maintenance of long-term specimen collections. Temperature should be monitored.

A-6.15 Quality assurance (assessment) and quality control

A laboratory quality management system is crucial to ensure correct testing results. For BBS, some of the important elements of a quality management system include documentation of procedures and records management, SOPs, external quality assessment (assurance), use of quality-control specimens, and quality-control retesting of a sample of survey specimens.

All laboratories should employ quality-assurance practices. Many laboratories participate in external quality assessment; that is, an evaluation of the testing process by an outside organization (e.g. a national reference laboratory). Such assessment can help to identify existing or potential problems, and activities include:

- site visits with direct observations, and review of testing procedures;
- quality control of survey test results by retesting a subset of specimens already tested (e.g. all positive and 5–10% of negative specimens retested by the national reference laboratory); and
- proficiency testing of all survey staff performing HIV or other testing.

Proficiency testing involves testing a panel of blinded samples sent to testers at testing sites and laboratories by a recognized proficiency testing programme provider. These samples are sent to the sites on a pre-defined frequency (2–3 times a year). Test results are submitted to the programme provider to be analysed for accuracy of testing. Poor performance is then addressed by retraining, increasing supervision or taking other corrective actions. Given the relatively short duration of

most BBS, survey sites that conduct testing should be monitored closely. Proficiency testing samples should be distributed monthly or perhaps every 6 weeks for the duration of the survey.

A-6.16 Supply chain considerations

The following considerations apply to supply management:

- the number of tests and consumables that need to be ordered should be determined.
- the amount of hardware equipment (e.g. testing instruments) needed per survey site should be determined.
- when procuring commodities, expiration dates should be considered. Some commodities, especially testing reagents, have a short shelf life (a few months) so they should not be ordered too early or in too large a quantity. Partial deliveries may need to be arranged with vendors or suppliers. Extra stock (10–20%) should be ordered to account for training, repeat testing or losses. To minimize wastage of perishable materials, they should not be ordered until all required approvals (not just ethical approvals) have been obtained.
- supplies that need to be ordered from abroad often have a long delivery time and may be subject to or delayed by customs.
- all key equipment (e.g. pipettes, enzyme-linked immunosorbent assay [ELISA] machines, refrigerators or freezers and incubators) should be well maintained and in good working condition.
- proper storage of supplies includes keeping them in a clean, organized locked shelf or cabinet and stored according to the manufacturer's instructions in a well-ventilated room away from direct sunlight. Supplies should be organized by expiration dates so that older supplies are used first (i.e. first expiry, first use).
- labels, printers and barcode readers should be procured, or the availability of preprinted labels ensured, before initiating the surveys.
- the United States Agency for International Development (USAID) Laboratory Logistics Handbook, which outlines how a supply chain should be designed and managed to ensure the availability of commodities, may be a useful resource *(11)*.

A-6.17 Biosafety and biosecurity

Biosafety and biosecurity measures are intended to assure the safety of participants, health-care workers and laboratory workers, and should be applied at all stages of laboratory activities, from specimen collection to laboratory testing and specimen storage. "Laboratory biosafety" describes the containment principles, technologies and practices that are implemented to prevent unintentional exposure to pathogens and toxins, or their accidental release *(12)*. "Laboratory biosecurity" describes the protection, control and accountability for valuable biological materials within laboratories that prevent their unauthorized access, loss, theft, misuse, diversion or intentional release *(13)*.

Laboratory managers and survey investigators are responsible for ensuring biosafety and biosecurity. The following should be available where specimens are handled (e.g. collected, processed, transported, tested and stored): protective gloves, laboratory coats, eye protection, waste bins for biohazardous materials, sharps containers, pipette tip disposal containers, regular waste containers, alcohol, bleach and other disinfectant solutions. Arrangements must be made for transport of biohazardous materials for subsequent incineration or autoclaving and re-use of suitable instruments. Access to post-exposure prophylaxis should be available in the event of accidental exposure to infectious agents (e.g. HIV) by survey staff.

A-6.18 References

1 Hauser RM, Weinstein M, Pool R, Cohen B. Conducting biosocial surveys: collecting, storing, accessing, and protecting biospecimens and biodata. National Academies Press. 2010 (http://www.nap.edu/catalog/12942.html, accessed 26 August 2016).

2 Dicker LW, Mosure DJ, Levine WC, Black CM, Berman SM. Impact of switching laboratory tests on reported trends in Chlamydia trachomatis infections. Am J Epidemiol. 2000;151(4):430–435 (http://www.ncbi.nlm.nih.gov/pubmed/10695602, accessed 26 August 2016).

3 Ghanem KG, Melendez JH, McNeil-Solis C, Giles JA, Yuenger J, Smith TD et al. Condom use and vaginal Y-chromosome detection: the specificity of a potential biomarker. Sex Transm Dis. 2007;34(8):620–623 (http://www.ncbi.nlm.nih.gov/pubmed/17308500, accessed 26 August 2016).

4 UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Guidelines for using HIV testing technologies in surveillance – 2009 update. Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2009 (http://apps.who.int/iris/bitstream/10665/164358/1/9789241598057_eng.pdf?ua=1, accessed 30 August 2016).

5 WHO. HIV assays operational characteristics: HIV rapid diagnostic tests (detection of HIV-1/2 antibodies): report 17. Geneva: World Health Organization (WHO); 2013 (http://www.who.int/hiv/pub/vct/rapid-diagnosis-tests-report17/en/, accessed 10 August 2016).

6 Centers for Disease Control and Prevention. Recommendations for the laboratory-based detection of Chlamydia trachomatis and Neisseria gonorrhoeae – 2014. MMWR Recomm Rep. 2014;63(RR-02):1–19 (http://www.ncbi.nlm.nih.gov/pubmed/24622331, accessed 30 August 2016).

7 WHO. Strategies and laboratory methods for strengthening surveillance of sexually transmitted infections. Geneva: World Health Organization (WHO); 2011 (http://apps.who.int/iris/bitstream/10665/75729/1/9789241504478_eng.pdf, accessed 30 August 2016).

8 WHO. Guidelines for the safe transport of infectious substances and diagnostic specimens. Geneva: World Health Organization (WHO); 1997 (http://www.who.int/csr/emc97_3.pdf, accessed 30 August 2016).

9 Dong Q, Nelson DE, Toh E, Diao L, Gao X, Fortenberry JD et al. The microbial communities in male first catch urine are highly similar to those in paired urethral swab specimens. PLoS One. 2011;6(5):e19709 (http://www.ncbi.nlm.nih.gov/pubmed/21603636, accessed 30 August 2016).

10 ARUP Laboratories. Sample collection for the diagnosis of STD using nucleic acid amplification tests: Chlamydia trachomatis, Neisseria gonorrhoeae, and Trichomonas vaginalis. Salt Lake City: ARUP Laboratories; 2014 (https://www.aruplab.com/Specimen-Handling/resources/pdf/SampleCollectionForSTD.pdf, accessed 30 August 2016).

11 United States Agency for International Development. Laboratory logistics handbook: A guide to designing and managing laboratory logistics systems. Arlington, VA: USAID; 2009.

12 WHO. Laboratory biosafety manual. World Health Organization (WHO). 2004.

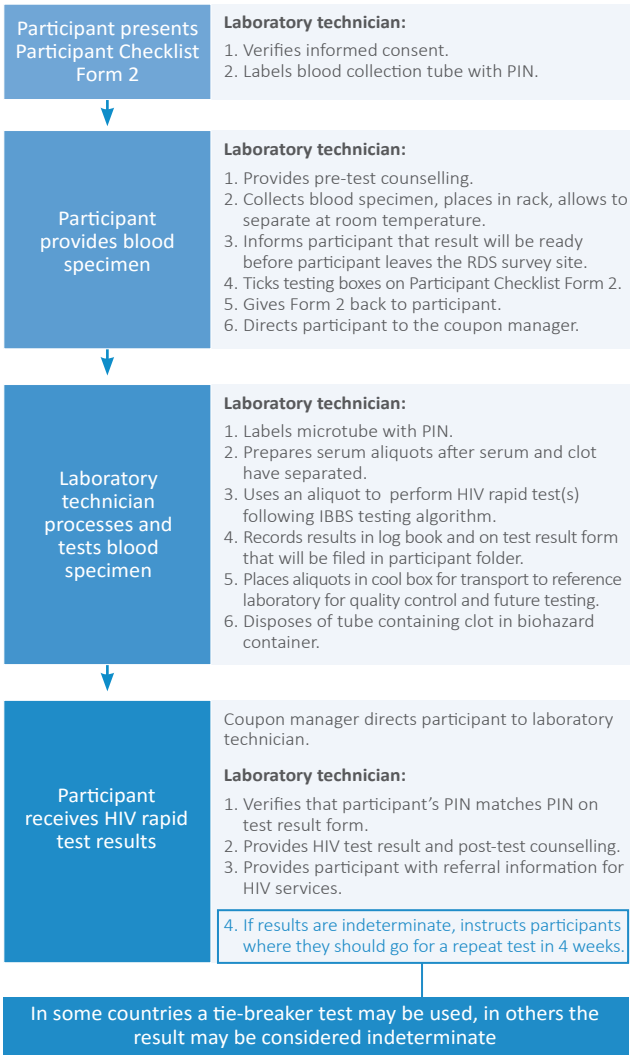
13 WHO. Biorisk management: Laboratory biosecurity guidance. Geneva, World Health Organization (WHO). 2006 (http://www.who.int/csr/resources/publications/biosafety/WHO_CDS_EPR_2006_6.pdf, accessed 30 August 2016).

Section A

7. Preparing biological data-collection instruments

This chapter focuses on the data instruments used for the collection, processing, testing, transport and storage of biological specimens. Procedures for each of these activities vary; for example, because of differences in survey protocols, choice of assays and algorithms, the site of specimen collection and testing, and whether specimens will be stored for future testing. Survey investigators should prepare a detailed flow chart to illustrate each step. Figure A-7.1 gives an example of a flow chart.

Figure A-7.1 Sample flow chart illustrating the process for rapid HIV testing at a survey site, ranging from obtaining consent to providing HIV test results



Key terms

Algorithm: A step-by-step procedure.

Biomarker: A biological characteristic measured through laboratory or other tests, generally using a biological specimen such as blood for testing; examples include HIV antibody or viral load.

Calibrate: To compare and adjust readings of an instrument with those of a standard; adjust experimental results to take external factors into account or allow comparison with other data (source: adapted from Oxford Dictionary).

Chlamydia trachomatis: Causative agent of chlamydia, a bacterial sexually transmitted infection.

Molecular testing: Testing for nucleic acid sequences (e.g. DNA or RNA) as opposed to immunologic testing for antigens or antibodies.

Neisseria gonorrhoeae: Causative agent of gonorrhoea, a bacterial sexually transmitted infection.

Optical density: Intensity of colour as measured by a spectrophotometer. In some enzyme immunoassays (EIAs) that detect HIV antibodies, more intense colour (higher optical density) generally indicates more antibody, although EIAs are designed to be qualitative, providing an answer of reactive or nonreactive.

A-7.1 Specimen tracking forms

Specimens are often collected at one site and transported to another for testing and storage. Paper or electronic forms are used to track specimens from the point at which they are collected through processing, testing, shipping, storage and final disposal. To describe specimen integrity, survey staff should maintain a record of the dates and times specimens were collected and the temperatures at which they are stored. This is especially important for tests that require specimens to be in good condition (e.g. not exposed to extreme temperatures or multiple freeze–thaw cycles). Testing standard operating procedures (SOPs) and test kit inserts include guidance for optimal specimen handling and storage processes.

Specimen tracking forms should include the following information:

- survey participant ID;
- specimen ID (if different from participant ID);
- date of specimen collection;
- survey site location;
- type of specimen;
- date specimen stored in cool box, refrigerator or freezer at site;
- storage temperature during transport – minimum and maximum – and transport conditions, (e.g. on dry ice, wet ice or cool packs, or at room temperature);
- date specimen transported to another site (e.g. off-site or referral laboratory) for testing or storage;
- name and location of second site (e.g. off-site or referral laboratory);
- dates and times at which the laboratory received, processed, tested and stored the specimen;
- number and volume of specimen aliquots or dried blood spot (DBS) cards or spots;
- physical location and temperature of storage (e.g. refrigerator no. 2 or freezer no. 3) so that specimens can be easily retrieved;
- initials or other form of identification for each person who handled the specimen at each step; and
- other information locally determined to be needed.

Appendix I-13 provides an example of a specimen tracking form.

A-7.2 Specimen transport logs

Specimen transport logs are used to record the movement of specimens from one site (e.g. the collection site) to another (e.g. the off-site or referral laboratory). This form accompanies the specimens. A copy should be made and kept at the site where other survey information is maintained.

Logs may include the following information:

- initials or other form of identification of person (e.g. survey manager or laboratory technician) who handed specimens over and person (e.g. driver) who received them;
- date and time specimens picked up by transport;
- temperature of cool boxes (if used); and
- date and time specimens delivered to laboratory or storage site.

Some of these data (e.g. time and temperature) are critical for some specimens and assays (e.g. urine for *Neisseria gonorrhoeae* or *Chlamydia trachomatis* molecular testing). For consistency, and to allow for testing of specimens in the future, temperature monitoring should be performed for all specimens, even those that can be transported at room temperature.

Appendix I-14 provides an example of a specimen transport driver log.

A-7.3 Refrigerator or freezer temperature logs

Temperature monitoring should be performed at the start of each workday at any location (survey site or off-site referral laboratory) where specimens or temperature-sensitive reagents are stored.

Temperature should be monitored using calibrated and certified laboratory thermometers. In spaces that experience larger temperature fluctuations (e.g. a cool box), temperatures should be monitored using special thermometers that record the minimum and maximum temperatures experienced in the container during the time specimens were being held.

Appendix I-16 provides examples of specimen temperature monitoring log.

A-7.4 Test results form

Just as survey questionnaires are used to record responses to questions, the test results form is used to record results of specimen testing for biomarkers. These testing data may be entered into a separate database from the survey interview database. The two files can be merged using the survey participant identification number (ID), to link biomarker test results with interview data.

Some laboratory testing machines automatically produce assay results and other information. This automation reduces transcription errors and allows survey staff to export testing data for merging with other data files (e.g. the interview data).

The test results form (electronic or paper) should include the following:

- survey participant ID number;
- assay name, lot number and expiration date;
- name or other form of identification of person performing each test;
- date testing performed;
- results from specimen testing for biomarkers;
  - optical density values, if applicable;
  - titres, if applicable;
  - individual assay and final test (algorithm) result;
  - test quality control and calibrator results, if applicable;
- comments;
  - notes about specimen integrity (e.g. gross hemolysis or bacterial overgrowth);
  - notes about specimen volume (e.g. amount of plasma received for testing, or quantity not sufficient);
- date result reported to survey site or survey investigators, if applicable (e.g. testing performed off-site); and
- date test result provided to participant, if applicable (e.g. on-site test results and same-day test result provision).

Appendix I-17 provides an example of a test results log.

Section A

8. Eligibility for participation

Participant eligibility for a BBS usually depends on age, recency of a risk behaviour, sex, and city of work or residence. Depending on the survey's objectives, potential participants may have to meet other criteria as well. This chapter defines the target population and shows how screening for eligibility yields the survey sample.

Key terms

Eligibility criteria: The description of all the requirements a potential survey participant must meet in order to participate.

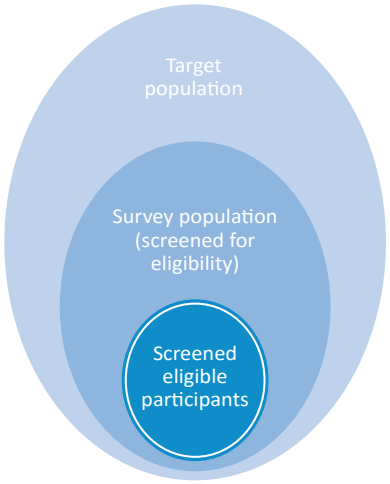
Eligibility screening: Asking potential participants a series of questions to determine survey eligibility.

A-8.1 Defining the population

The target population is the group of people about which investigators are interested in drawing conclusions. This population is generally described broadly – for example, "men who have sex with men (MSM) in Mexico City" – and is not specific enough to provide guidance about who should be "allowed" to participate in the survey. For example, the above definition may include MSM who are not eligible to participate in the survey, including those that are not currently sexually active or who are too young to provide informed consent.

The survey population is a subset of the target population narrowed to account for practical considerations related to sampling methodology and eligibility criteria. In a time-location sampling (TLS) survey, this population might include men who congregate in venues where MSM are known to socialize and who are randomly selected by survey staff. In a respondent-driven sampling (RDS) survey, it would be men who are identified as MSM by other MSM, who are referred by a previous participant, and have access to the survey site. The survey sample is the subset of the survey population that actually participates in the biobehavioural survey (BBS).

Figure A-8.1 The relationship between target population, survey population and eligible participants



An individual must meet the eligibility criteria for a survey in order to participate. For example, a BBS conducted among female sex workers (FSW) may define eligibility as aged 15 years or over who have exchanged sex for money or goods at a brothel in the past 6 months. For MSM, the definition may include men aged 15 years or over who have had anal sex with other men in the past 6 months. For people who inject drugs (PWID), eligibility may be defined as aged 15 years or over and having injected drugs within the past 30 days or the past 6 months. Eligibility criteria ideally should be precise, time bound, measurable and reliable.

"Precise" means that the information given by the potential participants can be used to accurately describe the survey participants. "Time bound" means that eligibility based on a certain behaviour is limited to a specific time frame, so that the potential participants can accurately recall and describe it.

"Measurable" means a criterion can be enumerated; for example, participant age in years. "Reliable" means the information provided by potential participants can be verified if needed. It can be difficult for eligibility criteria to fulfill these standards. Survey staff have to be well trained in order to spot potential participants who may not tell the truth.

A-8.2 Verifying eligibility

Eligibility is verified through an eligibility screening tool. Typically, participants are asked a series of questions to determine their eligibility. Other techniques may also be used; for example, PWID may be required to show injection track marks or demonstrate how to assemble a needle and syringe, and prepare drugs. Survey investigators should design a clear set of screening questions and a decision-algorithm so that the eligibility of potential participants is determined in a standardized fashion. In addition to excluding participants who do not meet the criteria, investigators may screen out potential participants who meet exclusion criteria; for example, those who are intoxicated (thus, unable to comprehend the questions or provide informed consent) or who have already participated in the survey.

A-8.3 Elements of eligibility

The following sections discuss considerations for elements of eligibility: age, sex, risk behaviour, time period, place of residence or work, and language.

Age

Typically, members of the target population must fall within a certain age range to be eligible. The minimum or maximum age for participation depends on the influence of age on the risk behaviour in question, and the national policy regarding the minimum age at which someone can provide informed consent. All BBS have a minimum age for this reason; however, few have an upper age limit. It is important to follow the national policy regarding minimum age for participation in a survey. In some places, investigators may be legally obliged to report to authorities if participants are found to be minors. If, during formative assessment, a large or particularly at-risk proportion of the target population is found to be below the national required minimum age for participation, investigators may need to obtain special permission from their parents or guardians for them to participate. Sometimes it may be possible to sample underage people without parental consent if they are independent (i.e. they obtain their own livelihood) or if getting such consent puts them at risk. Local laws and guidelines, and rules and regulations set by the local institutional review board (IRB) should be consulted. Chapter A-3 provides additional information on ethical considerations of survey conduct and the human subjects review process.

Sex

BBS are typically restricted to one sex, with the exception of surveys of PWID, which often include both sexes. For MSM and transgender person (TG) surveys, investigators need to weigh the merit and costs of combining these populations into one sample or keeping them separate, which often results in one population not being surveyed. Male and female sex workers (SW) may represent distinct HIV epidemics and thus are better sampled in separate surveys. For surveys that include both sexes, eligibility screening should include the sex of the potential participants so investigators can determine if each sex is proportionally represented in the final sample.

Risk behaviour

To be eligible to participate, potential participants must engage in the risk behaviour being studied. For example, a survey of PWID should not include noninjecting drug users. A survey of women who sell sex for money should not include those who only engage in transactional sex. If investigators are interested in sampling those who engage in transactional sex as well, the eligibility criteria should be altered accordingly. Similarly, a survey among MSM should not include men who identify as gay but have not had sex with another man.

During eligibility screening, potential participants should be asked questions that help verify their eligibility. For example, in a PWID survey, they may be asked to show track marks or to describe the injection process. Screening staff may probe MSM or SW for their knowledge of locations associated with the respective behaviour or terms used by the population; for example, venues they attend or terms they use to denote certain behaviours or groups of people.

Time period

Narrowing the participants to those who have engaged in specific behaviours within a specific time period allows investigators to focus on those who have engaged in risk behaviours most recently. Participants who have engaged in the risk behaviour recently will also be better able to describe the behaviour, and are less likely to be subject to recall bias. MSM who did not have sex with another man during the specific time period, or SW who did not exchange sex for money, goods or services during the specific time period, will not meet the eligibility criteria.

Place of residence or work

Most surveys have a defined sampling area, and therefore make residence or venue attendance (for selected venues) an eligibility criterion. For example, investigators may want to recruit only those who are residents of or work in the city/district/province. If

individuals in the population migrate often, investigators may define residence as having lived in the area for a certain time. Residency information may also help inform local programming needs after the survey, or allow for estimations of the population within the survey area. While those who do not live or work in the sampling area may contribute to the HIV epidemic, they are unlikely to access prevention, care and treatment services and consequently may not reflect local knowledge, attitudes or practices.

Language

Survey investigators should take into consideration the languages spoken in the region of the survey. Individuals should be proficient in the language(s) in which the survey is conducted in order to participate.

A-8.4 Exclusion criteria

Exclusion criteria are characteristics of potential participants that prohibit their enrolment. Such criteria may include being intoxicated (thus, unable to comprehend the questions or provide informed consent) or not presenting a valid coupon in an RDS survey.

9. Sampling strategy

Sampling is the process of selecting individuals from a source population to participate in a survey. The purpose of this is to generate information about the population without having to study the entire population, thereby reducing time and cost.

Once the target population has been defined, investigators have the option of selecting every member of the group, or selecting a subset of members (i.e. a sample). A sampling strategy describes how the sample will be selected. It also determines whether and how "statistical inferences" can be made from the sample to the source population; that is, the extent to which the sample is representative of the source population. An appropriate sampling strategy can generate unbiased estimates that:

- accurately describe the population;
- can be compared with estimates for other populations; and
- can be compared with estimates of the same population produced by other surveys or at different times.

This chapter briefly describes concepts used in sampling, types of sampling, and factors involved in selecting an appropriate sampling strategy.

A-9.1 Concepts used in sampling, and types of sampling

A-9.1.1 Concepts used in sampling

The main concepts used in sampling, discussed below, are:

- statistical inference
- sampling bias
- validity and reliability
- precision
- stratified sampling
- sampling domains

Statistical inference

Statistical inference is the process of using information from a sample to make inferences, or conclusions, about the population from which the sample was taken. If people are sampled in a biased manner, the sample may not be representative of the source population. Inferences or

Key terms

Convenience sampling:	The selection of individuals from a population based on accessibility and availability.
Location or venue:	Places where the target population spend time and may be accessed. Used interchangeably with "site".
Nonprobability sampling:	A method of sampling where the probability of sampling is unknown.
Reliability:	The degree to which a technique or approach provides consistent results if repeated multiple times.
Sampling:	The process of selecting individuals from a source population to participate in a survey.
Sampling bias:	The situation in which some population members are systematically more or less likely to be sampled than others, leading to biased estimates.
Sampling frame:	The source material or list from which a sample is drawn in order to represent the population.
Sampling interval (SI):	Defined as N/n, where N is the population size and n is the sample size.
Site:	A place where survey sampling occurs. It may be a clinic, office or apartment for RDS surveys, or a public space (e.g. a street corner, bus stop, brothel or club) for TLS surveys.
Statistical inference:	The process of using information from a sample to make inferences, or conclusions, about the population from which the sample was taken.
Validity:	The degree to which an estimate is likely to be true and free of bias (systematic error).

conclusions made on the basis of observations from such a sample may not be valid or reliable.

Sampling bias

Bias refers to the difference between an estimate and the "truth". Sampling bias occurs when some members of the population are more or less likely to be included in the sample than others. Without adjustments, the estimates produced by such a sample are not representative of the population.

There are two types of sampling error: systematic and nonsystematic. Nonsystematic sampling error, also known as random sampling error, is the error that results from taking a sample rather than measuring every individual in the population. For example, 100 different samples of the same population will each produce different estimates. The degree to which the estimates differ will depend on the size of the sample relative to the total population. Nonsystematic sampling error can be decreased by increasing the sample size. For example, if two surveys of the same population differ only in their sample size, the survey with a larger sample size will have less random sampling error than the survey with a smaller sample size.

Systematic sampling error is generally caused by either a poor sampling design or failure to follow sampling protocols. Unlike random (nonsystematic) sampling error, systematic sampling error cannot be predicted, calculated or accounted for. Examples of systematic sampling error include incomplete sampling frames, and samples that include people who do not meet the eligibility criteria.

A sampling frame is the source from which a sample is drawn. It is a list of all those within the source population who can be selected. When using simple random sampling, the list contains all individuals. When using cluster sampling, the list will include groupings of individuals known as clusters. Examples of clusters for target populations might include venues where sex is sold, and venues where men meet male sexual partners.

Some specific sampling (or selection) biases are especially relevant to biobehavioural surveys (BBS) and include the following:

- self-selection bias – Some people may have more motivation to participate in a survey if they have an interest in the subject matter or if the reimbursement for participation is very high.
- healthy worker bias – People who are healthier are more likely to be able to enrol in a survey than less healthy members of the population.
- differential referral – In surveys that ask participants to recruit others, individuals may not recruit randomly from the population. In chain-referral surveys, participants may recruit people who are more like them or who are more popular within their social network.

Validity and reliability

The term validity, as it is used in BBS, refers to the extent that the information collected in a survey answers the research question. Reliability refers to consistency; it is the extent to which a survey is likely to yield the same result when repeated. Whereas validity is concerned with the survey's success at measuring what the researchers intended to measure, reliability is concerned with the consistency of the actual measurement instrument or procedure.

There are two types of validity: internal validity and external validity. Internal validity refers to how valid the survey results are based on the survey design and any other factors that might influence accurate results. External validity refers to the populations, settings and variables to which this effect can be generalized.

Precision

Precision reflects the extent to which repeated measurements (e.g. surveys using the same methods) yield the same results. Precise results (i.e. estimates with small margins of error and narrow confidence intervals) may not necessarily be accurate (i.e. close to the truth). Instead, precision is an expression of variability, or the amount of imprecision.

Stratified sampling

Stratified sampling is the process of dividing members of the population into homogeneous subgroups before sampling. It is done to reduce variability or random sampling error, thereby improving the precision of the estimates produced by the sample. The strata should be mutually exclusive (i.e. every member of the population must be assigned to only one stratum) and exhaustive (i.e. every population member must be included). For example, if sex workers (SW) operate through three types of venues (brothel, street and residence), the sampling frame can be developed separately for each type of venue, and a predetermined proportion of respondents, reflecting the overall distribution of SW, can be drawn from each stratum. If the proportion from each stratum is not known, a similar effect is achieved by ordering the sampling frame by type of venue, unless there is an intent to oversample within some strata.

Sampling domains

When separate estimates are desired for different subsets of the population (e.g. different subtypes or different geographical locations), separate samples should be drawn from each. Aggregation of data from different sampling domains (e.g. to obtain national estimates from multiple local sampling domains) must be done with appropriate analytical techniques and weighting (Chapter C-1 provides further information on this).

A-9.1.2 Types of sampling

The two main approaches to sampling are nonprobability sampling and probability sampling. A nonprobability sample is one that produces estimates of unknown precision and reliability, and for which inferences should not be made or should be made with caution. A probability sample is one that produces estimates that can be inferred to the source population with a known level of precision and reliability.

Nonprobability sampling

Where populations are hidden because of stigma, discrimination criminalized behaviours or where sampling frames do not exist, nonprobability samples are often a necessity. Examples of nonprobability sampling methods include convenience, quota and snowball (a type of chain-referral) sampling. This section discusses where such types of sampling may be used. Further details of these methods are given in Section A-9.2.2.

Probability sampling

Probability sampling refers to sampling methods where every individual has a known (nonzero) chance of being selected, and the selection probability can be calculated. Examples of probability sampling methods include simple random sampling, systematic random sampling, stratified random sampling, cluster sampling, multistage sampling and respondent-driven sampling (RDS). Clusters (and cluster sampling) may be viewed in terms of:

- location – conventional cluster sampling (CCS); for example, a survey among prisoners;
- social ties – for example, RDS; or
- time and space – time-location sampling (TLS); for example, sampling of mobile individuals at selected locations during selected time slots.

Simple random sampling

Simple random sampling is the most basic type of sampling. Each person is chosen entirely by chance, and each person has an equal chance of being selected. Because the sampling probability is equal for everyone, simple random samples are considered self-weighted and no sampling weights need to be calculated. However, if the sampling frame is large or the population is spread out geographically, survey implementation can be impractical. Because complete sampling frames are generally not available for hidden or hard-to-reach populations, simple random sampling is not feasible for most BBS.

Systematic random sampling

Systematic random sampling is the selection of individuals at regular intervals from an ordered list of individuals in the population, beginning at a randomly selected point. The list could be ordered either alphabetically or by other criteria such as address or size. The main advantage of using systematic sampling over simple random sampling is its simplicity. For example, if the total population were 1000, a systematic random sample of 100 individuals from that population would involve observing every 10th person on a list.

Stratified random sampling

Stratified random sampling involves categorizing members of the population into mutually exclusive and collectively exhaustive groups, then choosing an independent simple random sample from each group, based on that group's proportion to the total population. Investigators can also oversample in some strata to ensure that enough people are selected to obtain stratum-specific estimates. For example, if there is a relatively small number of respondents in a particular stratum (e.g. females aged 15–24 years), taking a random sample from the entire population would not produce a sample large enough to make inferences about that subgroup.

Cluster sampling

Cluster sampling can be used when it is impractical or impossible to list all individual population members, but it is possible to list locations or clusters where they gather. By randomly selecting a subset of clusters (e.g. venues or sites) and sampling population members at those locations only, the task can be made more manageable and also less costly, because of substantially reduced travel and labour costs. One limitation of cluster sampling is that it requires larger sample sizes than simple or systematic random samples.

Multistage cluster sampling

Multistage cluster sampling involves multiple stages of random selection of sampling units. The clusters selected at the first stage are called primary sampling units (PSU), and at subsequent stages are called secondary sampling units, tertiary sampling units, and so on. Multistage cluster sampling is typically used when the population or the sampling area is large. At any stage, clusters may be selected randomly or with probability proportional to size (PPS), until the last stage, when simple random sampling of individuals is performed within each of the selected clusters.

Respondent-driven sampling

RDS combines snowball sampling (in which participants recruit other participants) with a mathematical model to compensate for the fact that participation is not random. RDS is based on the principles of chain-referral sampling: it relies on respondents at each wave to select the next wave of participants. RDS expands upon chain-referral sampling by introducing several innovations that minimize sampling error and account for some types of selection bias. These innovations allow for probability-based selection of participants based on the network within the target population. RDS is one of the most popular methods to sample hard-to-reach populations, such as key populations (e.g. men who have sex with men [MSM], people who

inject drugs [PWID] and SW). However, it makes several assumptions in order to fulfill the requirements needed to make population-based estimates (see Section 9.7 for a discussion of the assumptions).

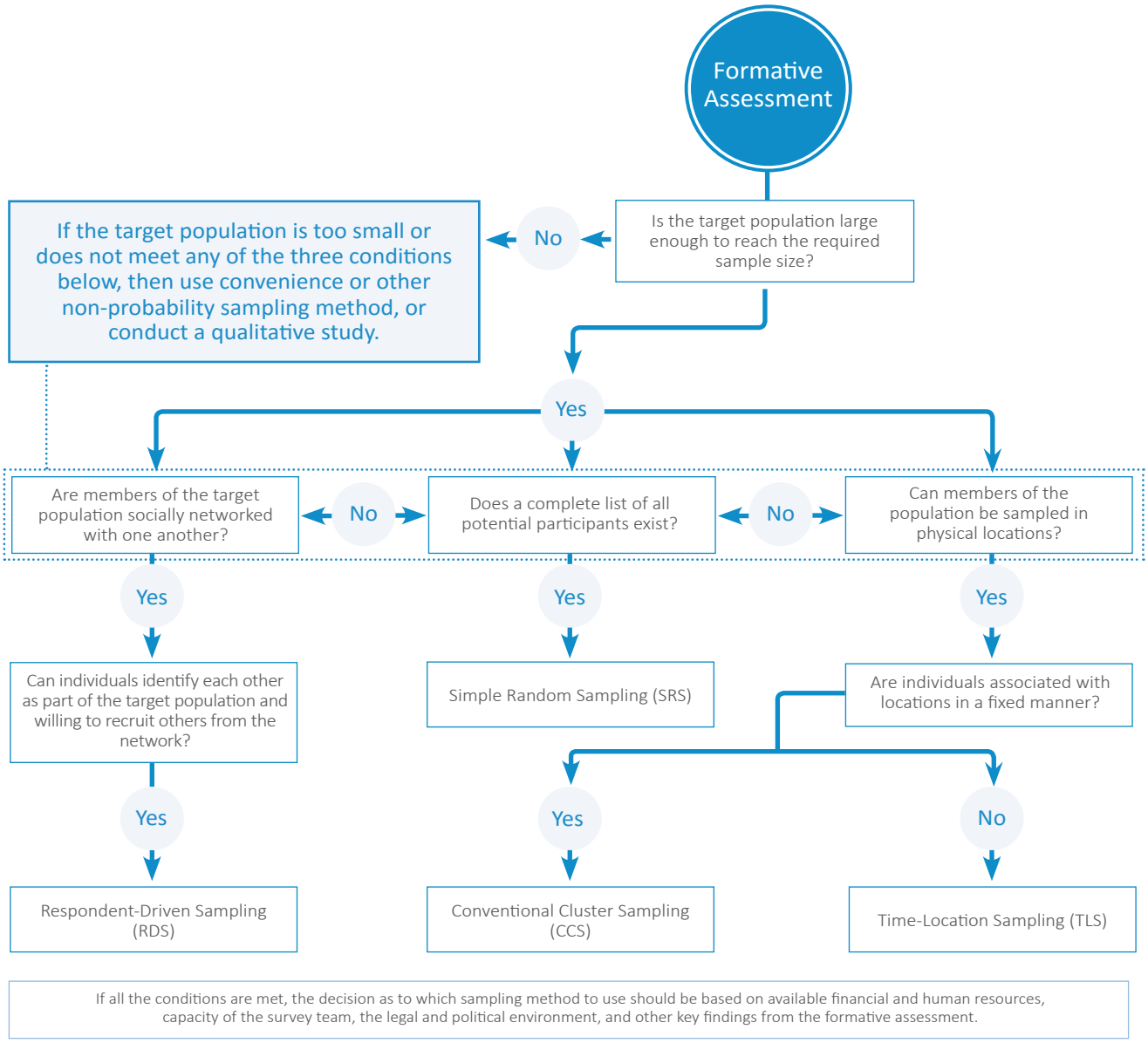
A-9.1.3 Selection of sampling strategies

Several factors must be considered before selecting a sampling strategy; for example, the existence of a sampling frame and the degree to which the target population is mobile, is networked, or can be identified and accessed at physical sites or venues. Figure A-9.1 shows a flow chart that can be used to facilitate the selection of a sampling strategy.

Factors to be considered in selecting a sampling strategy are as follows *(2-6)*:

- does a sampling frame exist? If a sampling frame exists, survey investigators can opt for a probability-based sampling design, such as random or cluster sampling.
- is the target population networked (socially connected)? If there is no sampling frame and it would not be possible to construct one, but formative research suggests that the target population is well networked, RDS could be used. For example, if PWID rely on other PWID to buy drugs, they may form a network, even if unintentionally. Similarly, networks may exist among SW and MSM.
- is the target population identifiable and accessible at physical sites or venues? If formative research suggests that a large proportion of the target population can be found at venues, and access to these venues is possible, then a sampling frame can be constructed and cluster-based sampling designs including CCS or TLS may be appropriate.

Figure A-9.1 Criteria for selecting a probability-based sampling method



A-9.2 Nonprobability sampling methods

Nonprobability sampling methods are flexible and easy to implement, but considerable care must be taken when using the information obtained from such samples because it is unlikely to truly represent the population being studied. Sampling bias is likely and, even if the direction of bias is known, the magnitude of bias is usually unknown. Usually, sampling bias also impedes comparison with other surveys. This section discusses when to use a nonprobability sampling method, types of these methods, their strengths and limitations, and ways to use them.

Key terms

Chain-referral sampling: A sampling technique where enrolled survey participants refer or recruit potential participants from among their acquaintances (peers).

Purposive sampling: A method of sampling used to select individuals with specific characteristics; for example, sex workers who take anti-retroviral treatment.

Quota sampling: A type of purposive sampling that specifies how many people with each characteristic will be sampled.

A-9.2.1 When to consider using a nonprobability sampling method

Nonprobability sampling methods are most useful when conducting qualitative research or when the probability sampling is either unimportant or impossible. Nonprobability sampling may be used when:

- the target population cannot be reached through a probability sampling method;
- because of the small size of the target population, it is impossible to achieve the sample size required for a probability method;
- for some other reason (e.g. resource constraints or low participation rates), it is impossible to attain the sample size required of a probability method, even if the target population is large;
- funds are limited;
- data must be obtained quickly; and
- the research capacity among investigators or staff is limited.

During survey planning, investigators should decide whether to use probability or nonprobability sampling methods. Probability-based methods are preferable but not always feasible. Sometimes, investigators begin with the intention of using probability sampling but are unable to generate a sampling frame and have no alternative but to adopt a nonprobability design. Nonprobability methods are vulnerable to bias; hence, it is important for investigators to identify possible biases in their BBS and assess how such biases affect the results.

A-9.2.2 Types of nonprobability sampling methods

There are many kinds of nonprobability sampling methods, the most common being convenience, purposive, quota and chain referral, as discussed below (this information complements the discussion about types of sampling in Section A-9.1.2, above).

Convenience sampling

In convenience sampling, investigators select participants who are most easily available and accessible. This is fast and less resource intensive than other methods. Investigators may visit a number of locations and ask people to participate at each location. Selection probabilities are unknown, and not everyone who meets the eligibility criteria is guaranteed selection. Convenience sampling can be useful, particularly when surveying a population that is extremely small. It is best used in the exploratory stage of a research and surveillance activity. An example of convenience sampling is recruitment of participants based on availability (e.g. PWID who congregate near a train station or patients at a health clinic).

Take-all approach

The take-all approach is a kind of convenience sampling in which investigators attempt to sample every person that meets the eligibility criteria; it is most commonly adopted when the target population is small. Even under those circumstances, however, it is unlikely that 100% of the target population will be sampled. The take-all approach can be used at the final stage of a multistage cluster sample when identified clusters have few eligible people, or if there is limited access to many clusters. In these instances, it is more efficient to sample all eligible people who are approached for participation.

Purposive sampling

Purposive sampling is the process of sampling individuals with specific characteristics. Whereas a convenience sample might enrol all SW who are easily identified,

a purposive sample might only select SW who have a regular partner. This method may be adopted when investigators only need information on a subset of a population, or are trying to ensure a diverse sample along certain specific dimensions.

Quota sampling

Quota sampling is a type of purposive sampling that specifies how many people with a given characteristic must be sampled; it is generally used to ensure that particular segments of the population are sampled. For example, in a survey of PWID, investigators could decide that they want to interview 80 PWID, comprising 20 female and 20 male homeless PWID, 20 people who also sell sex, and 20 additional PWID of any characteristics. This approach does not establish statistical criteria for selecting the individuals with these characteristics; hence, their selection is not probability based and the sample is not representative of the entire population of PWID.

Chain-referral sampling

Chain-referral sampling encompasses a group of sampling methods in which individuals recommend others to investigators or directly recruit others to participate in the survey. The sample grows through peer referral. Snowball sampling and indirect sampling are types of chain-referral sampling. Indirect sampling is discussed in Section A-12.4.

Snowball sampling

Snowball sampling gets its name from the image of a snowball that gets larger as it rolls down a hill, collecting more snow as it goes. Snowball sampling is a type of chain-referral sampling in which a few well-informed initial participants are selected purposively and then help to enlarge the group. Investigators ask them to provide names of other individuals who are well informed on the survey topic. These people are then invited to participate in the survey, and the process continues *(7)*.

In another version of snowball sampling, the initial well-informed individuals are asked to recruit others who are eligible for the survey. The recruits, in turn, are asked to recruit others, who then recruit others of their own, until the sample size is reached. As in the other nonprobability methods described above, the probability of selection of any given participant is unknown.

A-9.2.3 Strengths and limitations of nonprobability sampling methods

As with all sampling methods, nonprobability sampling methods have strengths and limitations, as indicated in Table A-9.1.

A-9.2.4 How to use a nonprobability sampling method

Surveys using nonprobability sampling methods are relatively easy to implement. In all other regards, they face the same requirements for implementation as probability sampling methods. For example, nonprobability sampling methods also require standard operating procedures (SOPs), informed consent and standardized questionnaires.

Surveys that use nonprobability sampling methods can be implemented in many different ways. For example, investigators can sample people and ask them to come to a survey site, or the survey can be administered wherever survey staff finds potential participants. The appropriate method may depend on the social visibility of the target population *(8)*. If investigators forgo the selection of a survey site, it is important to consider how to test people for HIV and other sexually transmitted infections (STIs) in public settings. This scenario is described in more detail in Section A-9.5 on CCS. If investigators adopt snowball sampling, they may need to create coupons to assist participants in recruiting peers and directing those peers to the survey site. Considerations for preparing coupons and peer recruitment can be found in Section A-9.7 on RDS.

Table A-9.1 Strengths and limitations of nonprobability sampling methods

Strengths	Limitations
• Relatively inexpensive • Relatively easy to implement • Can be conducted with any sample size • Can be used when there is no sampling frame	• Sample is not representative of the underlying population because the selection probability of participants is unknown. • Results from multiple surveys are not comparable. Nonprobability methods have great potential for sampling bias, so it is impossible to determine if observed differences in results are due to changes over time in behaviour, or merely in the sample. • No basis for assessing the precision or reliability of results.

A-9.3 Cluster sampling overview

The purpose of this section is to give an overview of cluster sampling methods used in BBS. It covers characteristics common to all cluster sampling techniques and the most important differences among the methods commonly used to survey key populations.

Cluster sampling is the method of choice in many surveys because it can provide quick, relatively inexpensive and representative data on populations. It has been adapted for use with key populations over the past 15–20 years.

Cluster sampling can be used when it is impractical or impossible to list all individual population members, but when it is possible to list locations (clusters) where target population members gather, and where they can be identified and approached for survey participation. This is often the case for key populations. The sampling frame is an exhaustive list of clusters. Clusters are places where population members live, work, socialize or conduct other activities related to their defining characteristics.

Cluster sampling, in its simplest form, involves a two-stage process. First, a subset of clusters are selected from an exhaustive list of all possible clusters; then individuals are selected from within each selected cluster. Surveys carried out in very large areas sometimes have additional stages beginning with selection of a subset of geographical areas at the first stage (e.g. districts, wards or census enumeration blocks of a large city), followed by selection of clusters within the selected geographical areas at the second stage, and the actual sampling of individuals within the selected clusters at the third stage.

A-9.3.1 Types of clusters

These guidelines distinguish between two main types of clusters: physical clusters and time-location clusters. Physical or fixed clusters are places where a fixed number of people is present (e.g. a brothel). Surveys selecting such clusters are using CCS. Time-location clusters are places where the number of people may change over time (e.g. people attending a bar or club). Surveys selecting such clusters are using TLS.

A-9.3.2 Strengths of cluster sampling

Cluster sampling can provide a statistically valid representative sample of a population spread over a large geographical area, without requiring a complete list of eligible population members for its sampling frame. The survey team instead samples a limited number of (randomly) selected clusters. By selecting clusters, and listing and sampling population members in those clusters only, the sampling task becomes more manageable and less costly compared to the alternative of simple random sampling.

A-9.3.3 Limitations of cluster sampling

People who gather together in clusters may be more likely to have certain traits in common with each other than they do with people in other clusters. For example, young SW might prefer to solicit clients on corners with other young SW whereas older SW may choose to solicit clients elsewhere.

Furthermore, people who inject heroin do not necessarily interact with people who inject methamphetamine. This social or behavioural clustering (also referred to as design effect) can skew survey results if not accounted for in analyses (see Chapter C-1). Accounting for the design effect in analysis usually has the effect of increasing the standard error, so results from cluster surveys tend to have less statistical precision than results from surveys that use simple random sampling and an equal sample size.

When cluster sampling is venue based, which is typically the case, it has the additional disadvantage of only capturing the subset of the population who frequent venues, and who are identifiable and accessible at those venues. So, population members who do not frequent venues are excluded, as are population members who frequent only inaccessible or unidentifiable sites. To the extent that these population members differ from those who are accessible at venues, venue-based cluster sampling may introduce bias.

Examples of key population members who may not be reached through venue-based sampling include:

- SW whose clients contact them through brokers or via mobile phone, and who do not solicit clients from physical locations;
- MSM who only gather in private locations such as homes or who find sexual partners only on the internet; and
- PWID who inject in private locations (e.g. homes), or who get their drugs and injecting equipment from friends or family rather than directly from dealers.

A-9.3.4 Planning for cluster sampling

A number of decisions should be made early on in the survey planning process to determine how the sampling frame will be constructed and what information will be required to complete it.

Ensure a venue-based cluster sample will be appropriate for the population

Be sure the definition of the target population is compatible with a survey design that uses cluster sampling. For example, a survey of brothel-based SW will not include SW who find their clients via phone or internet. If formative research or existing information indicates that a significant proportion of the target population will not be reached by cluster sampling, consider another approach.

Define the geographical catchment area

Typically, a survey will define a sampling area through political or geographical boundaries (e.g. a city's political boundaries). If it is believed that substantial urban and rural differences exist, then this may necessitate including enough respondents in both types of geographical areas and stratifying the analysis by urban and rural areas (see below).

Consider stratification

Stratification is the process of separating members of the target population into more homogeneous subgroups (or strata) before sampling. Stratification may become necessary if two or more subgroups of the target population differ so much from each other (e.g. in their risk behaviours, their burden of disease, or ethnicity) that it would be better to view them as separate populations. Common examples are brothel-based and street-based SW or MSM and transgender women, who may be so different from each other that investigators need to sample each subpopulation (stratum) independently. Prior knowledge about the cluster types and relative sizes of the subpopulations from formative assessments are used to assign every sampling unit (venue, time-location slot) in the sampling frame to one, and only one, stratum. Usually, stratification requires over-sampling (an increase in sample size) in some or all strata to allow for separate estimates for each stratum.

If stratification is not feasible because of budget or time constraints, it may be necessary to redefine the target population or venue types so that the survey is restricted to one subpopulation only.

Build design effect into sample size calculation

Simple random samples have a design effect of 1, but for cluster surveys, the design effect is usually greater than 1. The expected design effect can be used to guide how much to increase the sample size to compensate for the increased variance. Design effects are sometimes available from estimates of the primary outcome from previous surveys. If they are not, some speculation on the variance for important variables including the primary outcome may be required. Statistical software packages, including Epi Info, SPSS, SAS and Stata, can calculate the design effect for each variable in a given survey (see Chapter A-10).

A-9.4 Procedures for two-stage and multistage sampling

A-9.4.1 Two-stage cluster sampling

Most BBS using cluster sampling will use a two-stage sample design. PSUs or clusters are chosen at the first stage of sample selection, and individual respondents are chosen from within each of the selected PSUs at the second stage. This sampling scheme, adapted to meet the needs of the different subpopulations and local conditions, will likely satisfy the needs of most BBS efforts.

A-9.4.2 Multistage cluster sampling

If the sampling universe is very large, meaning the sampling area covers a large geographical area, additional sampling stages, such as province, district or segment of a city, can be added. For example, in a provincial survey of SW, investigators may include towns as the first sampling stage, venues as the second stage, and individual SW as the third stage.

A-9.4.3 Selecting primary sampling units

This section describes procedures for selecting sampling units at each stage of cluster sampling. Clusters selected at the first stage of sampling are called primary sampling units (PSUs), and clusters selected at subsequent stages are called secondary sampling units and tertiary sampling units. When two-stage cluster sampling is used, the PSU and the ultimate clusters from which individual respondents will be selected are the same.

Probability sampling requires that the selection probability of every respondent is known. With simple random sampling (SRS), every respondent has an equal probability of selection. With two-stage or multistage cluster sampling, the selection probabilities are the combined product of the sampling probabilities at each stage of selection.

Equal probability versus probability proportional to size
Sampling units (clusters) must be selected at every sampling stage. There are two ways to select clusters: by probability proportional to size (PPS) or by equal probability. When clusters are selected by PPS, it means they are selected proportional to their size, so larger clusters have a larger probability of selection. When clusters are selected by equal probability, it means they are selected without regard to size. So, large clusters have the same probability of selection as small clusters.

Key terms

Cumulative measure of size:	The cumulative sum (across clusters) of the estimated number of individuals who meet the approach criteria.
Estimated measure of size:	The number of individuals estimated to be present and eligible to be recruited during formative assessment or sampling frame development. This number helps investigators plan sampling strategies, and could be used to carry out selection of clusters using probability proportional to size (PPS).
Multiplicity:	The chance that any one sampling unit (person) could be captured in multiple clusters (i.e. an individual could be recruited in multiple venues).
Primary sampling unit (PSU):	The unit used for the first stage of sample selection (e.g. a brothel) – in a two-stage cluster sample, the PSU and the cluster are the same.
Probability proportional to size (PPS):	With PPS, a sampling unit (e.g. a cluster) gets selected according to its size (relative to that of others); hence, larger clusters have a larger probability of selection.
Random start (RS):	A randomly selected number between 1 and the sampling interval.
Sampling event:	The time period when field staff are at the venue identifying, selecting, recruiting and interviewing respondents.
Sampling unit:	In a statistical analysis, refers to one member of a set of entities being studied (e.g. bars or respondents).
Sampling universe:	The population to which the results can be inferred.

To use PPS selection procedures, the number of individuals associated with each PSU must be known in advance. This number, known as a measure of size (MoS), does not need to be an exact count. A rough approximation for each site may be good enough, and because it is rough, it is referred to as an estimated measure of size (EMoS).

Selecting clusters by PPS improves sampling efficiency by increasing the likelihood that larger clusters are selected. However, the larger the size of the cluster, the smaller the probability of each individual within that cluster being selected. To compensate for this, a weighting process is used in the analysis, which works by assigning

a weight to individuals from every cluster that is inversely proportional to the size of the cluster.

Self-weighted designs

When the combination of sampling probabilities at each stage produces a selection probability that is the same for individuals from every cluster, this is known as a self-weighted design. Combinations that will produce self-weighted samples include:

- PPS at the first stage and a fixed (equal) number at the second stage; and
- equal probability sampling (EPS) at the first stage and select all, or select fixed proportion (e.g. 20% from each cluster) at the second stage.

Statistically, the first of these (PPS with equal number at second stage) is the more efficient two-stage sampling design. However, it does require accurate measures of size prior to selecting clusters. The second (EPS with select all or fixed proportion at second stage) can be challenging to implement if very large clusters are selected. The final sample size also becomes unpredictable. As an alternative, it is possible to select fewer respondents (or an equal number of respondents) from clusters selected by EPS. However, the final sample may not be self-weighted.

Self-weighted samples are easier to analyse, but if the EMoS are very different from the actual measures of size (AMOS), then the sample may no longer be self-weighted. Therefore, it is always best to collect and record all information needed to conduct a weighted analysis, in the event that it becomes necessary to perform a weighted analysis. Note that even for self-weighted survey datasets the design variables are still needed in data analysis to correctly calculate standard errors and p-values.

Decision criteria for EPS and PPS

Ideally, investigators should obtain a current list of clusters with their EMoS, to facilitate PPS sampling. Choose EPS only if information on EMoS is not available. If information on EMoS is available and clusters have little or no variation in EMoS, investigators may choose EPS as well as it would equal a self-weighted design. However, in most cases EMoS will vary substantially across clusters, and will warrant a PPS design (Figure. A-9.2). If investigators choose to generate a self-weighted sample, and EMoS is available, PPS is the method of choice. For discussion on how to approach PPS vs EPS for TLS designs, see Section A-9.6).

A-9.4.4 Procedures for selecting sampling units

Determine how many clusters are needed and what the ideal cluster size should be

The number of clusters to be selected must first be determined before selecting clusters by PPS or EPS. The number of clusters will be a function of the desired cluster size (number of respondents to be sampled from each cluster), and the overall sample size for the survey. More information on how to sample clusters using EPS and PPS can be found in Appendix I-18.

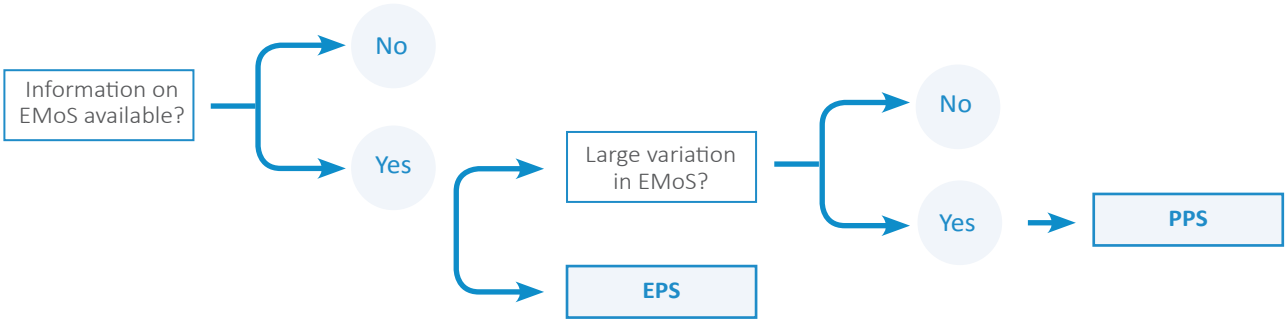
Desired cluster size

Sampled clusters should be small enough to make a listing of individuals feasible, but large enough to support a minimum number of sampled individuals per cluster. Generally speaking, a sample with smaller clusters will result in a smaller design effect than a sample with fewer larger clusters. But more clusters may also mean higher travel/transport costs. Another factor to consider is the degree of expected homogeneity (similarity) within clusters as opposed to between clusters, with respect to the outcomes of interest (e.g. behavioural variables, HIV status). If homogeneity is expected to be low, then fewer clusters will be needed.

Number of clusters to select

Survey statisticians often recommend that there be no fewer than 30 clusters (see also Chapter A-10). A common approach to deciding how many clusters to select is to divide the total sample size for the survey by the desired cluster size. The desired cluster size should be smaller

Figure A-9.2 Decision criteria for EPS and PPS



EMoS, estimated measure of size; EPS, equal probability sampling; PPS, probability proportional to size

than the (average) EMoS; that is, the expected number of eligible people in a given cluster (which should be available from the sampling frame). It is possible to sample a fraction of all eligible persons at a given cluster; however, cluster sizes should still be as large as or larger than the average EMoS in order to avoid shortfalls in obtaining the desired sample. Nonresponse (refusal to participate) should also be factored into the design. Some designs deal with nonresponse by inflating the initial sample size, others by doing replacement (see Chapter A-10).

Example of calculating number of clusters

1 - calculate the average EMoS from the sampling frame. For this example, assume EMoS = 10.
2 - only aim to sample a fraction of the average EMoS; for example, 30% to 50%. It is unrealistic to assume everyone in a given cluster can be sampled. In this example, investigators aim at a cluster size of no more than 3 or 5 (30–50% of 10). This improves efficiency of the cluster sampling (reducing design effect) and minimizes the risk of shortfall due to nonresponse.
3 - the required number of clusters, therefore, should be 600 (target sample size) divided by 3, 4 or 5; that is, 200, 150 or 120 clusters.

Collecting information to calculate sampling weights

Regardless of whether EPS or PPS is used, it is critical to document the information that will be needed to calculate sampling weights.

This involves the following steps:

- preserve the sampling frames with information about each cluster, including cluster name, cluster identification number (ID), MoS (if applicable) and selected units at each sampling stage.
- record sampling event information:
  - actual measure of size
  - # approached and eligible
  - # consented to survey
  - # nonresponse (e.g. refusals, incomplete interviews or incomplete blood samples).
- record individual information and answers to the multiplicity question (if applicable). Multiplicity refers to the possibility that some people may be present at different clusters at different times and so may get approached or sampled more than once.

Maintain the original sampling frame for reference, or in case a second round of cluster selection is required. Document the sampling information in accordance with the steps above; this information is required to calculate sampling probabilities.

A-9.5 Conventional cluster sampling versus time-location cluster sampling

A-9.5.1 Features of CCS and TLS

There are two major types of venue-based multistage cluster sampling typically used with key populations: conventional cluster sampling (CCS) and time-location sampling (TLS). Table A-9.2 summarizes the basic features of CCS and TLS.

Table A-9.2 Features of CCS and TLS

Conventional cluster sampling	Time-location sampling
CCS should be used when there is a relatively fixed population associated with fixed venues (sites) such that: • individuals are associated with only one venue • the same individuals are associated with the same venue no matter when it is visited • a complete list of individuals who are associated with the venue may be obtained or created Examples of situations where CCS might be appropriate include sampling of stationary: • brothel-based SW • prisoners • military personnel • PWID in residential treatment facilities.	TLS should be used when the target population is mobile, meaning different people may be present at a given venue at different times. It may be used when individuals come and go freely from the venue (site), and when their composition may differ depending on the time the venue is visited. In TLS: • individuals may be associated with more than one venue • it may be possible to create a list of attendees at a venue, but only of those who are present during the sampling event (time period) PSU in TLS contains both a time and a location element. The same physical location may be included in the sampling frame multiple times, to account for the variation in target population members frequenting the site at different times, hence the name "time-location" or "venue-time" (or "venue-date-time") cluster sampling. Examples of situations where TLS might be appropriate include: • SW at street-based sites, bars or nightclubs • MSM at bars or MSM cruising sites • PWIDs at shooting galleries or other injecting venues

A-9.5.2 Types of venues used in cluster sampling for key populations

Table A-9.3 on the next page lists the types of venues typically used for key populations (SW, MSM and PWID).

Target population	Conventional Cluster Sampling	Time-location Sampling
Sex workers (SW)	Brothels,[a] massage parlours,[a] hair salons	Brothels,[a] massage parlours,[a] bars, restaurants, discos, city blocks, streets, public parks
Men who have sex with men (MSM)		Bars, nightclubs, streets, parks, restaurants, bathhouses, railway platforms, beaches, social organizations
Transgender persons (TG)	Community of hijras linked to a particular guru (in India)[b]	Bars, nightclubs, streets, parks, restaurants, bathhouses, railway platforms, social organizations
People who inject drugs (PWID)	Drug treatment centres	Street locations, shooting galleries, injection parlours, syringe exchange centres
Young people	Households, schools	Workplaces, locations where young people gather
Long-distance truck drivers	Transport companies (with rosters of employees)	Truck stops, depots, border checkpoints
Prisoners	Prisons, jails, detention facilities	
Students	Schools	
Military personnel	Barracks, camps	
Hospital patients	Hospitals	
Miners	Mining camps	

CCS, conventional cluster sampling; TLS, time-location sampling

Note: Some of the examples are listed in both categories because the nature of the relationship between the cluster and the population may vary depending on the context.

[a] Brothels and massage parlours can be sampled using CCS or TLS. If the SW live or always work at the same brothel, CCS may be appropriate.

[b] Hijras is a term used for transsexual or transgender people in India who sometimes live in communities led by a guru.

A-9.5.3 Conventional cluster sampling

This section describes conventional cluster sampling (CCS) and the general steps required to prepare for using this method. CCS is used when population members are fixed in relation to the venues from which they will be sampled.

A-9.5.4 Overview of CCS design

CCS is a cluster sampling method. Typically there are two stages of sampling, but there can be more. The first stage of sampling consists of selection of PSUs, which can be drawn with equal probability or with PPS. At the second stage, the target population members in the selected clusters are randomly or systematically selected from the complete list of individuals in the PSU.

When to use CCS

Use CCS when there is a relatively fixed population associated with a site or cluster such that:

- individuals are associated with only one site or cluster;
- the same individuals are associated with the site or cluster no matter when it is visited; and
- a complete list of individuals who are associated with the site or cluster can be obtained or created.

The PSU in a two-stage cluster sample using CCS is the site or cluster. Examples of situations where CCS might be appropriate include sampling of students in schools, detainees in prisons, military personnel in barracks, PWID in treatment facilities and SW living in brothels.

Strengths of CCS

- CCS needs only an exhaustive list of clusters where eligible participants can be accessed and recruited.
- CCS is a robust sampling technique that has been commonly and successfully used in the past.
- under ideal circumstances, it is possible to create a self-weighted sample.

Limitations of CCS

- if the design effect is greater than 1, a larger sample size will be required to obtain a comparable level of precision to that of a simple random sample.
- obtaining self-weighted samples may be difficult if accurate measures of cluster size are not available in advance of the survey.
- although listing eligible respondents at venues is theoretically possible, it may be difficult to obtain a complete list.
- it may be necessary for the team to return multiple times to the site to find all the people who have been selected. If the team takes shortcuts and replaces respondents who are unavailable during the first visit, with those who are more easily available, this can introduce bias.
- though the members of the community are theoretically available at any time, the lists may change frequently. Prisoners, PWID in treatment facilities and military personnel, for example, may transition in and out of that location at any time.

CCS plan

A number of preparations specific to sampling in a site with a fixed population are required before the first sampling event can take place (see Appendix I-1).

- sample size – Determine the sample size using the methods described in Chapter A-10.
- method for sampling PSU – Two approaches exist for sampling clusters:
  - equal probability should be used when there is minimal variation in EMoS between clusters or if there is little information on the EMoS of each cluster; and
  - PPS should be used when there are fairly good EMoS for all the clusters in the sampling frame and high variation in MoS between clusters.

Individuals within clusters are sampled randomly or systematically. Two options are available for sample selection from each selected cluster (see Section A-9.4.4 on desired cluster size):

- option 1 – The sample size within each cluster may be *equal* (i.e. the same number of individuals are selected from each cluster). Choosing this option after selecting PPS at the first stage will produce a self-weighted sample, provided that the EMoS is accurate. This option also has the advantage of providing control over the final sample size.
- option 2 – The sample size within each cluster may be a *fixed proportion* of individuals from the total in each cluster. This option has the advantages of not requiring EMoS ahead of time, and it will produce a self-weighted sample. However, the disadvantage is lack of control over the final sample size.

Sampling key populations

Some special considerations are warranted for sampling key populations such as SW, MSM and PWID:

- community advisory board – Establish a community advisory board (CAB), made up of target population members. The board will provide a liaison to the community to facilitate trust and safeguard the community's interests. A terms of reference for the CAB can be helpful.
- team members – It is advisable to include key population members on the survey team, either as community consultants, or in some cases as interviewers or field team leaders. This can help with establishing trust and rapport with the community, and also with ensuring the identity of key population members.
- management permission – Locations or venue owners or managers, where relevant, may be approached for permission to conduct sampling at the site. If they refuse, some information about the location or venue should be collected (numbers and types of target population members) to provide some measure of possible sampling bias.

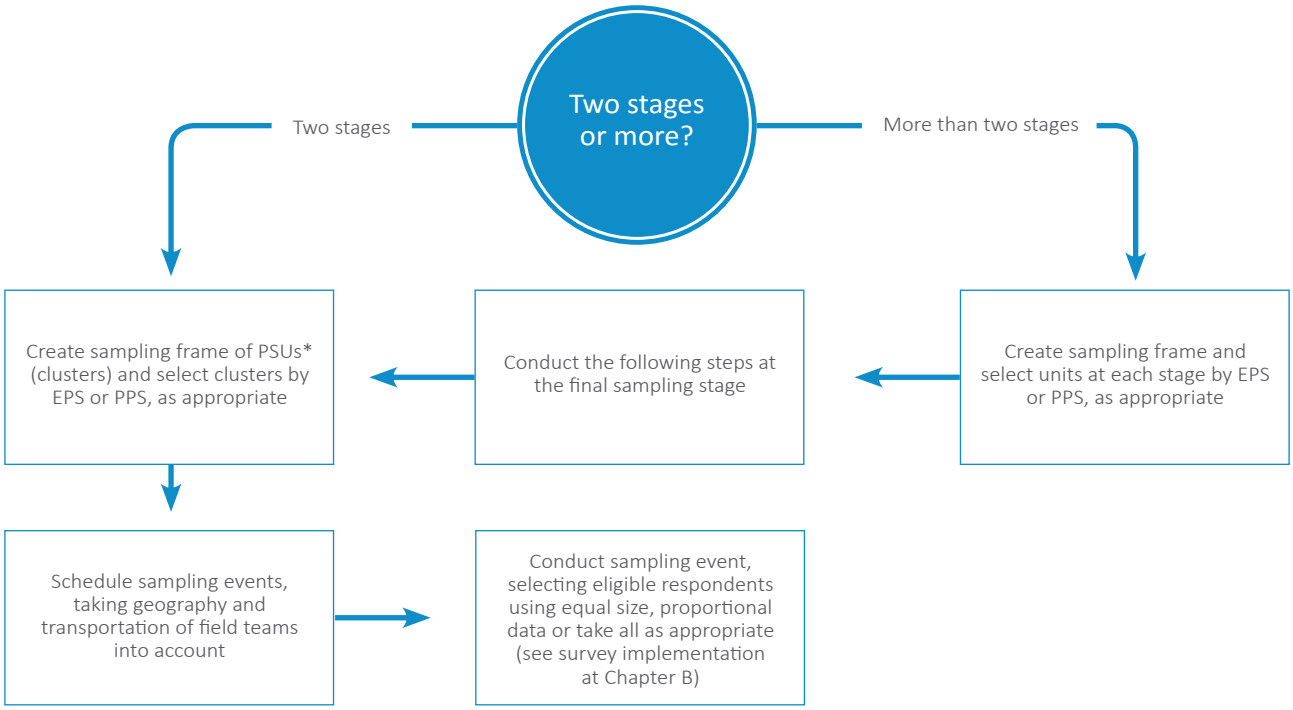
A9.5.5 Steps for using CCS

The steps for using a CCS are described here and can be found in the flow chart at Figure A-9.3.

1. eligibility criteria: Define the population being surveyed clearly, and also the types of venues where they will be accessed (i.e. inclusion and exclusion criteria). For example, in a survey of SW, the eligibility criteria might be women who work in brothels and are aged 15 years or over. This may be appropriate if there is a clear programmatic decision to focus on brothel SW. But it may exclude many women who work outside of brothels (e.g. those who are street based or bar based, or operate only through phone or social media sites).
2. catchment area: Define the geographical area that will be represented by the survey, and where sampling will occur. This might be an area defined by a political boundary like city limits (e.g. the district of Kampala) or a geographical boundary like "east side of the river".

3. sampling frame: The sampling frame is a comprehensive list of locations where the population can be found, and from which the sample will be selected. Sampling frame development can be easy or complex, depending on whether up-to-date lists of venues exists, or whether they need to be created. If existing lists are used, they must be updated to verify that they are current. That means locations where population members no longer gather must be removed from the list and new locations must be added. Typical types of locations that serve as clusters are brothels, lodges, nightclubs and street corners (for SW), schools or classrooms (for in-school young people), and barracks (for uniformed services). However, ultimately the clusters must match with whatever is specified in the eligibility criteria. If EMoS are available for clusters, they can be used to facilitate the sampling scheme.
4. sampling of clusters: If EMoS are available, clusters can be selected by PPS. When PPS sampling is used, larger clusters have a higher probability of selection than smaller clusters. This can make sampling more efficient; however, adjustments must then be made to account for the lower selection probability of people in smaller clusters. If EMoS are not available, clusters can be selected with equal probability. The number of PSUs selected will be a function of the sample size and the desired cluster size. This is described in more detail in Section A-9.4.3.
5. inclusion of sampled clusters: All selected PSUs must be visited and contribute to the sample even if the target sample size has already been reached.
6. sampling of individuals: Each PSU is expected to contribute survey respondents, according to the chosen option.

Figure A-9.3 Conventional cluster sampling steps



* Depending on the target population and the type of cluster, this may be as simple as obtaining a list of locations from a minor partner (e.g. in case of prisons) or as complicated mapping location such as brothels or sites where sex workers solicit clients (e.g. in the case of sex workers)

A-9.6 Time-location sampling

Key terms

Focus group: Semistructured group conversations conducted with usually 6–8 members of the target population under the guidance of a facilitator using an interview guide. Focus groups are convened to confirm or refute observations made in key informant interviews, and to gather additional information to facilitate survey preparation and implementation.

Key informant (interview): People who have first-hand knowledge of the target population or subject matter.

Time-location sampling (TLS): A cluster sampling technique that not only considers the venue (location) but also the time of sampling. TLS is used when the number of people attending a venue may vary over time or day. Examples include sampling of patrons at bars, clubs or bathhouses.

Venue (location) universe: The entirety of all venues from which clusters or time-location clusters are being drawn.

A-9.6.1 Overview of time-location sampling design

This chapter describes the steps involved in sampling participants using time-location sampling (TLS). TLS is used when participants are mobile with respect to the venues from which they will be sampled. For example, street-based SW or MSM attending clubs or bars may move around between venues, and will not be associated with any one venue in a "fixed" or stable manner. This is in contrast to CCS where sampling happens at conventional clusters or "sites" with "fixed" or stable populations (e.g. children in school or inmates in prisons).

To implement TLS as a cluster sampling method, a sampling frame is needed that allows all members of the population to have a nonzero chance of selection into the sample, and a method for calculating selection probabilities. The sampling frame should be constructed in such a way that it can capture the different types of people who may frequent venues at different times of the day and on different days of the week. This is accomplished by constructing a sampling frame that is composed not only of venues frequented by target population members, but also of time slots when the venues are operational. These combined venue or time clusters are known as time-location clusters.

An example of a time-location cluster might be a specific gay bar between 10 pm and 2 am on a Friday night. Men who are found at this time-location cluster may differ from men found at this same location at another time (e.g. 2 pm to 6 pm on Mondays), not only as individuals, but in terms of their levels and types of risk behaviours. For example, men frequenting sites on weekdays may be more likely to be unemployed and therefore have less disposable income. For this reason, it is important to be able to select venues during different time periods.

In its simplest form, TLS is a two-stage cluster sampling design. The PSUs are the time-location clusters. TLS clusters are selected by equal probability or PPS (in the same manner described in Section A-9.4.3 on selecting PSUs). The second stage comprises the sampling of population members at the selected time-location clusters. Because individuals can come and go from the venues, they must be sampled during defined sampling events. This is in contrast to CCS, where respondents are "stationary" with respect to the venue, and so can be listed and sampled at any time. To use PPS sampling with TLS, the "expected" EMoS at different times must be known in advance, and it is understood that by definition, the MoS will vary, so the estimated sizes can only serve as expected values.

A-9.6.2 Strengths of time-location sampling

TLS provides a means of sampling mobile populations in a representative and probabilistic manner.

A-9.6.3 Limitations of time-location sampling

- if the design effect is greater than 1, a larger sample size will be required to obtain a comparable level of precision to that of a simple random sample.
- the yield from individual sampling events can be unpredictable because of changeable MoS.
- likewise, obtaining self-weighted samples may be difficult if accurate measures of cluster size are not available in advance of the survey.
- in theory, recruitment of respondents should be evenly distributed throughout the sampling time interval. This can be difficult to manage logistically.

Venue universe versus time-location sampling frame

The location or "venue universe" is a list of all of the potential locations or venues in a given geographical area, including those that may not be appropriate for the survey. It contains information about the location or venue type and each of the potential day-time periods, EMoS of each time-location slot, any stratification information, and any information that informs the investigator if the venue and day-time period is appropriate to include in the sampling frame.

The "time-location sampling frame" is an exhaustive list containing all the relevant time-location clusters to be sampled from. The variables included are location or venue ID, location or venue name, time period and EMoS. Additional variables may need to be included, depending on factors such as target population and cultural context.

A-9.6.4 Time-Location Sampling Plan

These steps (see also Figure A-9.4) can be followed for developing a sampling plan for a TLS survey (*9*):

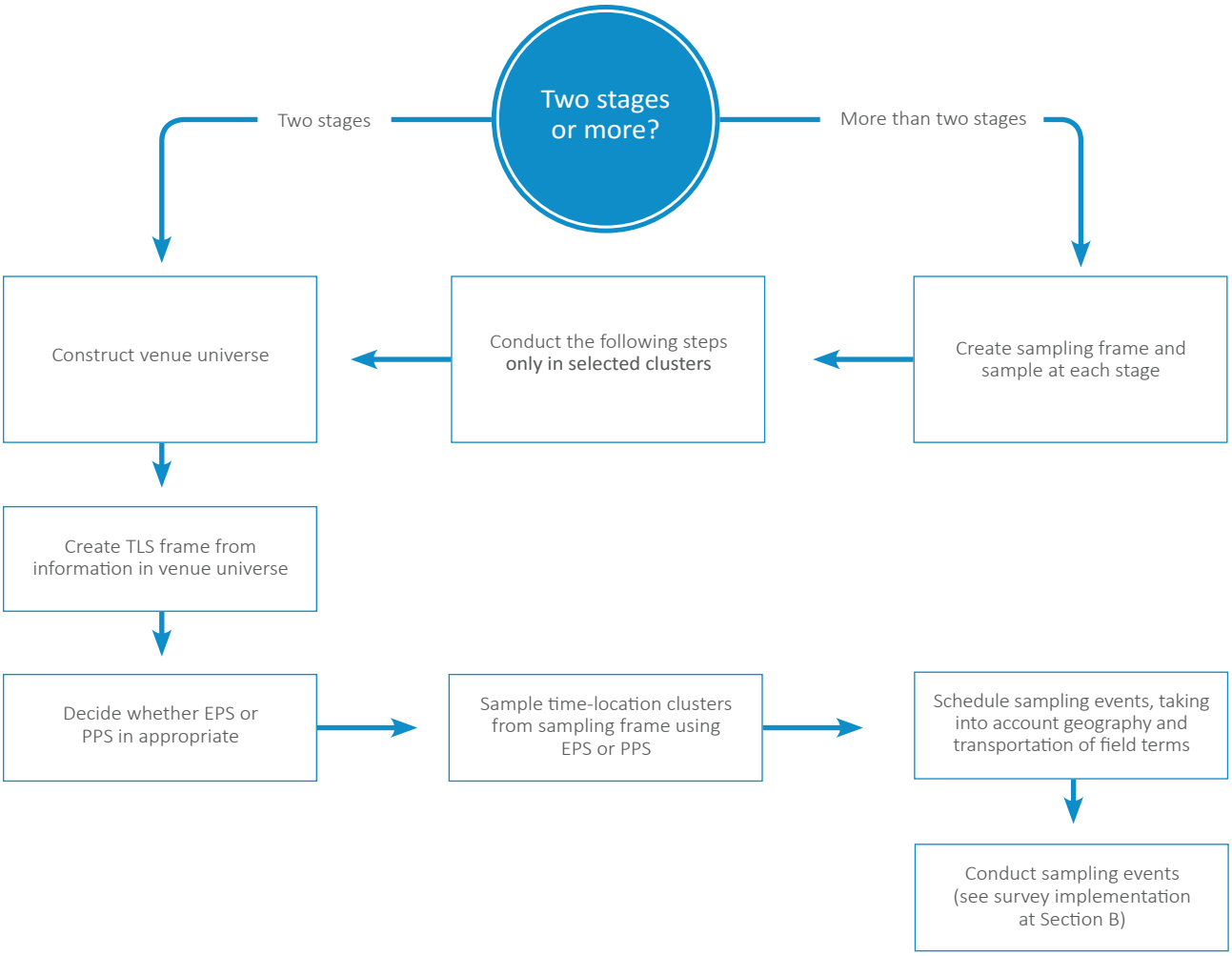
1. define eligibility criteria – Define survey eligibility criteria, keeping in mind that eligibility will have to be established quickly in the field and from visual assessments.
2. establish a CAB – Establish a CAB made up of target population members. The board will provide a liaison to the community to facilitate trust and to safeguard the community's interests. A terms of reference for the CAB can be helpful.
3. plan the team composition – Plan to include some key population members on the team as community consultants (liaisons) or interviewers/field team leaders to help with trust and rapport.
4. develop a sampling frame – Develop a comprehensive list of potential time-location clusters (sampling frame). This involves visiting venues and completing site information sheets that document information such as name of the site, name of contact person at site, physical location, characteristics and boundaries of site, type of site (e.g. for FSW this might include brothels, bars or street corners), estimated number of target population members expected to be found at the sites on different days and at different times, and frequency of target population members visiting sites. This last item is to provide some indication about how often venues should be included in the sampling frame. If target population members frequent sites daily, it would not be desirable to list the site in the sampling frame multiple times because of the risk of duplicate sampling. But if there is high turnover at a site, then the site can be included in the sampling frame multiple times.

   Some sampling frame development exercises record on the site information sheet peak times and times with few people at the site, and list maximum and minimum numbers of target population members expected at different days and times.

   At the end of the sampling frame development exercise, investigators can develop a comprehensive sampling frame consisting of all time-location clusters. The sampling frame should exclude time-location clusters that are likely to be too small (i.e. have too few attendees) to warrant a sampling event; venues and locations for which venue management refuses to grant permission for sampling events; and venues that are deemed too dangerous for survey staff.
5. select the clusters – Using the complete sampling frame, time-location clusters are selected (first sampling stage). These clusters may be selected either through EPS or PPS (see Section A-9.4.4 on procedures for selecting sampling units). The selected time-location clusters are then added to the "sampling event calendar" (i.e. a calendar table with cells for each day). Care must be taken to avoid days when sampling will be difficult or impossible, such as when staff may not work or certain venues may be closed or low in attendee numbers (e.g. public holidays or gay pride festivals).

Once the calendar is completed, preparations to conduct the field work can begin and field recruitment and enrolment initiated.

Figure A-9.4 Time-location sampling steps



EPS, equal probability sampling; PPS, probability proportional to size; TLS, time-location sampling

Time-location sampling considerations

This section presents considerations relevant for development of a protocol and SOPs. A protocol checklist (see Appendix I-1) and the TLS section of the survey design checklist (see Appendix I-4) can facilitate planning.

Formative assessment

By the end of formative assessment, investigators should have a good knowledge of the types of locations or venues where the target population can be accessed. The formative assessment may reveal location or venue types that investigators had not anticipated. It is important that all location or venue types be represented in the venue sampling frame.

General sampling event recruitment planning

- because public locations or venues are often relatively unstructured and each one is different, it is best to create strategies for recruitment in a variety of situations so that the field team can approach and recruit people in a systematic and efficient manner rather than according to a strict set of procedures that may not always fit well.
- geographical boundaries for each location or venue should be fixed and agreed upon by the team when the locations or venues are mapped.
- a standard text for introducing recruiters and the survey should be developed and rehearsed for each sampling event.
- recruitment should be done in small teams of no fewer than two people. The approach should ensure safety of the team while facilitating adherence to the SOPs.

Documentation of sampling procedures

All information relevant to each sampling stage should be documented and retained. This includes sampling frames at each stage (including time-location units at the final sampling stage), procedures for selecting sampling units at each stage, cluster sizes, number sampled in each cluster and nonresponse. Reasons for nonresponse may

include refusal of venue managers to allow for a sampling event to take place, and refusal or failure of an individual respondent to give consent or to successfully complete all stages of data collection.

Completeness of sampling frame
Create clear SOPs to collect an exhaustive list of locations and venues where individuals could be accessed for an interview and then develop a comprehensive sampling frame consisting of all the relevant time-location clusters to ensure that all locations or venues and day-time periods are listed. Ideally, the same methods can be repeated in subsequent surveys. It is not sufficient to simply copy the list of venues used in a previous survey.

Sources of venue information
Sources of venue information include key informant interviews and focus group members, websites, newspapers and field observations.

The location or venue data-collection form or site information sheet is the data instrument to record all information about each location or venue. The data collected through this form informs the location or venue universe and the time-location cluster sampling frame. The information collected is used to decide if the location or venue is appropriate for the survey, what day-time periods the target population is present and which venue-day-time periods should be listed on the sampling frame, as well as logistical considerations.

Potential information to be collected about each venue needed to construct the time-location sampling frame is listed in Table A-9.4, and examples of forms to capture the information collected are in Appendices I-19 and I-20. The data needed to execute a TLS survey differs by target population, culture and location, and may be guided by the formative assessment findings or observing some locations or venues themselves.

Missing location or venue types
The formative assessment may reveal that there are locations or venues where the target population gathers, but where it is impossible or unsafe to approach, recruit and interview them. For example, it may be unsafe (e.g. abandoned railway platforms) or impossible (e.g. brothel owner refuses to allow entry of survey staff) to conduct surveys at specific locations. Investigators should explore what subset of the population might be missed at these locations or venues and the potential sampling bias that may be introduced because these locations or venues are not in the sampling frame. This information should be reviewed when writing survey results to better understand survey limitations.

Location or venue suitability
This refers to criteria that a location or venue must meet in order to be included in the sampling frame. A location or venue is generally suitable if there are day-time periods with potential participants; if interviewing and HIV testing is logistically feasible at or near the venue; and where safety for staff and participants can be assured.

Location or venue attendance patterns
The SOP for construction of the location or venue universe and location or venue sampling plan should outline how to determine if attendance patterns are different at different time periods. If the attendance pattern differs significantly over different times of the day or different days of the week, or if the characteristics of the venue attendees belonging to target population change over time, the same location or venue may be included in the sampling frame multiple times based on changes in the attendance patterns over time/day.

Multiplicity
Multiplicity refers to the possibility that individuals may attend multiple locations or venues during the sampling period and therefore have a higher sampling probability. The most practical way to account for multiplicity is to ask each survey participant how often they attend other venues. This information can be used to calculate weights at the individual level to account for the frequency with which individuals attend venues.

Approach criteria
Approach criteria are used by the survey staff to decide who should be approached (and counted) for recruitment into the survey. These criteria are not identical with survey eligibility criteria; approach criteria allow for a decision about whether to count and approach an individual based on observation only. For example, in a survey of FSW, staff should approach only women who appear to meet the minimum age for participation. All survey recruiters need to use the same approach criteria, to avoid bias. The approach criteria are used to quickly gauge who is likely to meet survey eligibility criteria since both target population members and noneligible individuals may be present. The approach criteria should be informed by the formative assessment and venue observation (see Section B-1.3 for more information on approach criteria).

Method for selecting potential participants to approach at a sampling event
There are several ways to systematically sample potential participants at the selected time-location cluster (see Section B-1.3).

Selecting a method beforehand can help reduce sampling bias because some potential participants may appear more approachable than others. Rehearse the method before using it in the survey. It is possible to have more than one approach in a survey period, but it should not change during a sampling event.

Community liaison
It may be helpful to enlist the help of a community liaison to identify the potential participants who could be approached and recruited at a selected conventional

or time-location cluster. This person can be a member of the community being surveyed and who could help in establishing rapport with the community.

Length of sampling events
Sampling events can vary to between 2 and 4 hours. Sampling events typically are of equal length, but may differ in order to allow for more recruitment at low volume clusters, or less recruitment at high volume clusters. If a venue is active for a long time, for instance 6 pm to 2 am, it is better to break the time into two separate periods, or venue-day-time events, to avoid survey teams having to spend excessive amounts of time at a venue once the required number of people have been sampled.

Management permission
Location or venue owners or managers, where relevant, may be approached for permission to conduct sampling events in the future. If they refuse, some information about the location or venue should be collected to provide some measure of possible sampling bias. This information includes numbers and types of target population members.

Method for obtaining the actual measure of size
At each sampling event, the individuals who are eligible to be screened or recruitment (i.e. who meet the approach criteria) are counted (i.e. enumerated) and this number is recorded by a "counter". For any time-location cluster, the counter will be present for the entire time duration (even if interviews are completed before the end of the time slot) and will count each and every eligible and attending individual.

Plan for interviewing or testing outside the location or venue
Sometimes it is not possible to find an interview place or site at the selected location or venue that is appropriate to conduct interviews and testing. The protocol should account for this by either allowing for alternate-location or other-day interviews. It is fairly common to approach, screen and recruit people at a location or venue and then ask them to come to a different location for interview and biological testing.

- *alternate-location interviews:* When selected locations or venues cannot provide privacy, a tent or van nearby may be used, and participants escorted from the venue to the nearby survey location (it should be within walking distance).
- *other-day interviews:* Appointments on other days can be made if it is impossible to conduct interviews and testing at the time of enrolment. In this case, recruiters provide a survey ID and an appointment card with a time, date, and location for the interview. Use of this approach should be limited because some recruited persons may not come to the interview, leading to higher nonresponse. It may also complicate confidentiality. Still, it may be preferable to accommodate deferred data collection for some recruits or venues rather than losing data or having biased sampling.

Steps to determine the venue universe
Determining the venue universe essentially means to create a sampling frame that will then be used to select the clusters for inclusion in the survey. The following steps are discussed below:

- develop an SOP on how to create the sampling frame (describing the "venue universe");
- identify and describe the locations or venues;
- observe the locations or venues; and
- review and finalize the venue universe description.

Step 1: Develop an SOP for assessing the location or venue universe

- data elements: Decide on the information required for each location or venue.
  - number of attendees and time slots: What information is needed to decide on the time periods to include on the sampling frame?
  - venue attendance pattern: What information is required to determine if the venue is attended by people with different levels of risk at different times?
  - multiplicity: What information is required to determine if individuals at the venue also attend other venues?
  - safety and logistics: Where should data collection and testing take place?
- interviewees: Decide whom observers should talk to and what topics they should address.
- update the venue observation form as appropriate (Appendix I-20).

Step 2: Identify and describe the locations or venues

- data sources: Identify key informants or form focus groups. Key informants (single interviewees) should be knowledgeable about potential venues; they may be identified from among venue patrons, outreach service providers, venue managers or anyone else with good knowledge about the relevant venue "scene". After some key informant interviews, focus group discussions may be conducted to confirm or refute observations made in key informant interviews and gather additional information.
  - begin by asking key informants and focus group participants to identify locations or venues where the target population gathers and add them to the location or venue universe.

- find additional venues on the internet, in publications geared towards the target population, from outreach organizations and from survey staff.

- the location or venue universe should include all venues. Locations or venues that are not deemed "suitable" will be eliminated from the location or venue sampling frames.Determine venue type: Assign a code to each venue *type* that will be part of the venue code (e.g. bars = B, street corner = S).
- assign a venue ID: Use the location or venue type designation in combination with a number to create a unique venue ID (e.g. S3). Consider using a standardized code for province or geography if the survey is being conducted in multiple areas.
- add the venue's name and geographical information (address or street corner location) so that staff can visit the venue for observation.

Step 3: Observe the locations or venues

Start with locations or venues that are less understood and move to those that are more familiar. Use the venue observation form (at Appendix I-20) to document observations at each venue.

- establish rapport with the venue owner or manager; talk to venue staff and attendees.
- determine venue suitability using the venue suitability criteria.
- identify day-time periods appropriate for sampling events.
- gather other logistical information useful for field staff.

Step 4: Review and finalize the venue universe information

Review the gathered venue universe information for completeness and accuracy. Determine venues and day-time periods suitable for inclusion in the time-location cluster sampling frame (See Chapter B for more information about construction of the sampling frame).

Table A-9.4 Potential information to be collected about each location or venue (see Appendix I-20: Venue observation forms)

Venue information	Time periods	Target population observation	Venue logistics
• Venue type • Venue name • Venue code • Venue site contact info • Address of venue • Comments/description	• Minimum and maximum number of individuals affiliated with a particular location or venue time slot • Operational days of the venue • Peak and nonpeak days (days of the week when the maximum/minimum number of individuals are likely to be found) and EMoS on those days • Peak and nonpeak times (times of the day when the maximum/minimum number of individuals are found) and EMoS at those times	• Potential language barriers • Stability of location or venue • Visibility of location or venue • Observed behaviours • Observations about demographic characteristics • Seasonal variations • Mobility within as well as across location(s) • Detailed sketch of venue layout • Venue boundaries • Indicate intersections, landmarks to highlight venue boundaries • Location where key informants were spoken to • Locations where target population are grouped (if applicable) • If location or venue is a room in a building, be specific about how to find the room • If location or venue is a stretch of a street, show the side(s) of the street • Participant flow: use arrows to indicate the direction of entry to the site	• Safety concerns? • Potential interview/ testing location outside venue needed? • Barriers to recruiting? • Barriers to interviewing? • Barriers to testing? • Opened/closed since initial venue listing creation? • Parking for staff or van? • Comments • Space for interviewing, testing specimens at or around this site • Name and address of place • Nearest landmark • Contact person • Number of rooms available • Condition of rooms • Toilet available to collect urine samples? • Running water available?

The items on this list are provided as an example. Investigators should modify the information needs as appropriate.

The location or venue information is confidential data, and should be treated as such. Access to these records should be kept to a minimum number of people, records should be kept in a locked cabinet or encrypted files, and nonessential information should be destroyed when the data collection is over.

EMoS, estimated measure of size

A-9.6.5 First stage sampling: time-location

This section describes the sampling of time-location slots, including creation of the sampling frame from the venue universe and scheduling recruitment events.

Sampling frame construction and time-location cluster selection

Step 1: Create the sampling frame in database form from the information in the venue universe.

The "time-location cluster sampling frame" is a database with one line (row, record) for each time-location to be sampled from. The time periods should be chosen so that they are long (or short) enough to yield a sufficient (but manageable) number of sampled participants. Within this time slot, the patrons frequenting the venue should not change (e.g. patrons frequenting a bar early evening may differ from those later in the evening). The EMoS (i.e. the likely or anticipated number of eligible venue patrons or attendees) can be derived as part of the formative assessment or through venue observation, and should be generated with as much care as possible. The variables included are location or venue ID, location/venue name, time period and EMoS.

Step 2: Decide whether EPS or PPS is appropriate.

Regardless of the approach chosen, save the selected clusters with their unique cluster IDs on a separate worksheet for planning fieldwork. Maintain the original sampling frame for reference and in case a second round of cluster selection is required. Document the sampling information in accordance with the steps above: this information is required to calculate sampling probabilities and nonresponse.

Step 3: Select clusters with EPS or PPS.

Step 4: Create a calendar (see Figure A-9.6) for each field team.

- step 4.1: Ascertain staff availability; determine the dates and times the field staff will NOT be available to conduct recruitment events because of holidays, vacations or other planned absences. These dates and times should be blocked off the calendar.
- step 4.2: Block off other dates and times on the calendar when events cannot occur. Also consider staff burnout; in some settings, field teams are given a week off after 4 weeks of field work to return to their families.

Figure A-9.5 Sampling frame

Order #	Location (name, address)	Time	Estimated measure of size
1	Venue 1	9 pm–11 pm	8
2	Venue 1	11 pm–midnight	15
3	Venue 1	Midnight–2 am	10
4	Venue 2	8 pm–10 pm	5
5	Venue 2	10 pm–11 pm	15
...	Venue ...	...	...
18	Venue 8	8 pm–10 pm	14

Figure A-9.6 Calendar example with dates and times when sampling cannot take place blocked out

January 2017						
Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1 8–10 PM: S1 11 PM–1 AM: B2	**2** 7–9 PM: BR2	**3** 8–10 PM: BR1 11 PM–1 AM: S2	**4** 10 PM–Midnight: S3	**5** 7–9 PM: B1 10 PM–1 AM: BR3	**6** 6–8 PM: BR4 9 PM–Midnight: B3	**7** 7–10 PM: S4 11 PM–1 AM: B3
8	**9**	**10**	**11**	**12**	**13**	**14**

Step 5: Plan and schedule sampling events.

Schedule the sampling event for each of the selected time-location clusters. The specifics on how to do this should be determined in the local context and consider factors such as field team size, geography and infrastructure/roads. It may be helpful to order the selected venues geographically; for example, from north to south or near a main highway.

A-9.6.6 Eligibility and response rates in conventional and time-location cluster sampling: summary of numerical characteristics

Table A-9.5 highlights the information needed to facilitate the calculation of sampling probabilities at each stage and the response rate. There are a number of factors that influence the selection probabilities in cluster sampling, especially time-location sampling.

Table A-9.5 Data requirements for calculating sampling probabilities and response rate

Row	Characteristic	Conventional Cluster Sampling	Time-location Sampling
Approach			
A	# Selected individuals	Randomly or systematically selected from a list of eligible individuals	Systematically selected and approached from individuals who meet the *approach criteria* in a selected time- location cluster
B	# Refused approach	The selected person refuses approach	The selected person refuses approach
C	# Not present	The selected person is not present when the survey team is there[a]	n/a
D	# Accepted approach (A-B)	The selected person agrees to hear about the survey	The selected person agrees to hear about the survey and be screened for eligibility
E	Acceptance rate (#Accepted/#Approached, D/A)		

Row	Characteristic	Conventional Cluster Sampling	Time-location Sampling
Eligibility screening			
F	# Age or language ineligible	n/a The list provided to or created by the survey team ideally includes only eligible participants[b]	
G	# Who live outside geographical catchment area		
H	# Previous participants[c]		
I	# Who were not the correct sex/ gender		
J	# Who had not engaged in risk behaviour during the time period		
K	# Ineligible using any one of the eligibility criteria[c]		
L	# Eligible for participation (E-K)		
M	% of those approached who were eligible for participation (L/D)		
Participation			
N	# Who refused to participate after eligibility screening or consent		
O	# Who consented to participation in survey		
P	# Who consented to provide a specimen for HIV testing		
Q	Survey participation rate (O/D)	Normally, in CCS, only eligible individuals will or should be in the sampling frame and therefore offered participation, whereas in TLS that is not always possible. Individuals who are ineligible should not be considered when computing the participation rate.	
R	HIV testing participation rate (P/D)		
S	# Survey records lost[d]		
T	# Incomplete survey records[e]		
U	# Complete records available for analysis		
V	Enrolment rate		

[a] Schedule up to two return trips to interview those who are not present, if possible.

[b] If a large proportion of individuals interviewed appear to be ineligible, review the eligibility criteria with site stakeholder, re-create the list and resample.

[c] Approached individual may tell the recruiter they have already participated, or they may be identified by another one of the field staff. This may happen when the reimbursement for participation is high.

[d] Occasionally, paper survey forms are inadvertently destroyed or the electronic record is lost during transfer.

[e] Some analysts may choose to include incomplete records or records of those who consent to the survey but not the HIV testing in their analysis. This number is meant to represent the number of records with complete information.

A-9.7 Respondent-driven sampling

A-9.7.1 Overview of respondent-driven sampling

Respondent-driven sampling (RDS) is a peer-driven chain-referral sampling method that was first implemented in 1994 by Douglas Heckathorn *(10)*. This probability-based sampling methodology is useful for sampling populations that lack a sampling frame. It relies on participants' ability and willingness to identify and recruit other members of the target population into the survey. In RDS, recruitment relies on peers recruiting their peers rather than survey investigators selecting the sample.

RDS surveys start with a small number participants called "seeds" who are recruited by investigators or survey staff. These seeds participate in the survey and are then asked to invite peers within their social network to participate in the survey. A social network can include friends, colleagues and other personal contacts who know one another and who share characteristics that define them as members of the target population. Thus, for a survey among PWID, investigators enrol a small number of PWID seeds who then invite PWID they know to join the survey. These participants then recruit their peers, who are also invited to recruit other PWID who have not already participated in the survey. This process continues until the sample size is reached.

Differences between RDS and other chain-referral methods

Chain-referral sampling comprises a group of sampling methods that begin with a convenience sample of initial subjects (seeds) through which wave 1 subjects are recruited; these wave 1 subjects in turn recruit wave 2 subjects, and so on *(11)*. Snowball sampling is perhaps the best known chain-referral sampling method. RDS is similar to snowball sampling in that the target population must be socially networked so that participants can invite their peers into the survey. Although snowball sampling is useful for rapidly identifying potential participants, the methodology is biased. For instance, participants may recruit peers with characteristics similar to their own, producing a sample that is not representative of the target population. Additionally, if participants are allowed to refer an unlimited number of peers, people with larger social networks will dominate the sample. These limitations prevent investigators from drawing inferences from the sample and generalizing them to the entire target population *(12)*. In snowball sampling, investigators neither need to record who referred whom or estimate participants' personal network sizes, whereas in RDS they do.

Key terms

Convergence: The point at which the sample characteristics no longer change, no matter how many more individuals enter into the sample. Convergence is an indication of seed dependence, like equilibrium, but is based on the population estimate for a given variable. Whereas equilibrium is based on the wave, convergence is based on the order of enrolment into the survey.

Coupon: An invitation – by participants to peers – to enrol in an RDS survey.

Equilibrium: The point at which the distribution of participant characteristics is similar between waves
Page 128 of 296
(1). Equilibrium, like convergence, is an indication of seed dependence with respect to RDS, but is calculated based on the sample.

Personal network size (or degree): Number of personal contacts or peers who are part of the target population. Such contacts may include friends, acquaintances, partners, coworkers and relatives.

Primary compensation: Also called primary incentive, this is money or an item given after the first visit is completed to compensate for time and transportation.

Recruitment chain: All participants sampled through the same seed and connected through the resulting waves constitute a recruitment chain.

Secondary compensation: Also called secondary incentive, this is money or an item given at the second visit for each recruited peer that participates in the survey. Compensation for transportation may also be provided at the second visit.

Seed: A participant that is recruited by investigators or survey staff. All RDS studies begin with the selection of at least one seed.

Social network: A social structure made up of a network of personal contacts who share characteristics that define them as members of the target population.

Wave: The set of participants in a given number of recruitments from a seed. Individuals in wave 1 are those recruited directly by the seed. Individuals in wave 2 are recruited by those in wave 1.

RDS minimizes biases associated with chain-referral sampling by limiting the number of peers that each participant can recruit, which prevents people with larger networks from dominating the sample. It requires longer recruitment chains (i.e. more waves), but these should eventually result in a sample that is no longer biased by the characteristics of the purposefully selected seeds.

RDS is both a sampling method and an analytical method. RDS analysis considers that participants have different probabilities of being recruited and uses statistical weights based on the participant's personal network size or degree, and their recruitment patterns to estimate the prevalence of variables of interest. A participant's personal network size reflects the number of peers that could have recruited the person into the survey.

When to use RDS

RDS can only be used if the target population is socially networked and its members can recognize and recruit one another. All subpopulations and subgroups within the network must be connected. For example, to conduct an RDS survey among FSW, the subpopulation who sell sex in brothels should have at least some ties to the subpopulation who sell sex on the street. If this is not the case, investigators may want to conduct two separate RDS surveys: one among brothel-based SW and another among street-based SW. Social networks may also differ by other things such as neighbourhoods, age, ethnicity or nationality.

Because RDS recruitment is peer driven, it is often used when members of the target population do not congregate in large numbers in public venues. It is also employed when stigma and discrimination towards the target population make other sampling methods difficult. However, RDS is not only used among hard-to-find populations. Groups recruited using RDS include university students *(13)*, migrant populations *(14)*, and heterosexual men *(15)* and women at high risk of HIV *(16)*.

Strengths and limitations of RDS

Investigators choose RDS as their sampling method for numerous reasons; for example:

- it allows for sampling of less visible segments of the target population;
- the target population conducts recruitment;
- the sample can include individuals who attend venues and those who do not;
- it maintains or increases the privacy of the target population;
- it facilitates field operations and minimizes logistical needs;
- it minimizes the number of additional questions needed to create sampling weights; and
- it allows for computing population-based estimates.

On the other hand, limitations of RDS include the following:

- RDS relies on several assumptions that must be met in order to produce valid results, and it can be difficult to measure some of these assumptions, especially the assumption of random recruitment within a network;
- some social network components may be separated from others by factors such as geography, language or age, which results in a bottleneck in which recruitment occurs within one part of the network that is not linked to other parts;
- it is difficult or impossible to confirm the validity of RDS-based survey estimates;
- analysis of RDS-based data is challenging; and
- it is difficult to measure nonresponse in recruitment (e.g. when peers are offered a coupon but choose not to join the survey).

RDS theory: functional and analytical assumptions

Before implementing an RDS survey, investigators should understand what makes RDS a probability-based sampling methodology, and the assumptions upon which it relies. RDS involves two types of assumptions: functional and analytical *(17)*.

Functional assumptions

Functional assumptions of RDS are that:

- respondents know one another as members of the target population;
- respondents are linked by a network composed of a single component; and
- the target population is large enough to allow sampling with replacement.

An RDS survey will fail or be compromised if these functional assumptions are not fulfilled or are violated substantially. These assumptions, which should be examined during the formative assessment before RDS protocol development, are discussed below.

Respondents know one another as members of the target population

RDS assumes that members of the target population can recruit other eligible members to be in the survey. Therefore, RDS is feasible only when individuals can recognize one another as part of the target population. For example, MSM may recognize one another because they have sex with and socialize with other MSM. However, RDS may not be successful in situations where most SW work in isolation (e.g. by finding clients on the internet or working in private homes).

Respondents are linked by a network composed of a single component

Members of the target population must be part of a larger social network where each individual is directly

or indirectly socially connected to all other individuals through a series of social connections. These connections must be close enough to facilitate recruitment; in other words, individuals must interact with one another often enough so that a sufficient number of people can be recruited in the time available. Participants should know multiple people who meet the survey's eligibility criteria.

Participants should be able to recruit both people with whom they have weaker relationships, such as acquaintances, and people with whom they have stronger ones, such as friends and family. Bottlenecks occur when recruitment becomes "trapped" within a certain subgroup of the survey population. Members of the subgroup are socially linked to other members of the survey population; however, recruitment becomes bottlenecked because they disproportionately recruit other members of the same subgroup. Bottlenecks reduce the precision of RDS estimates because the bottlenecked group is more homogenous than members of the larger survey population (*18*).

The target population is large enough to allow sampling with replacement

Sampling with replacement means that each participant can be sampled more than once. With a fixed population, a person's chance of being selected for participation would be the same at any stage of the sampling process. In reality, sampling in RDS occurs without replacement *(19)*. Assuming that sampling is with replacement allows investigators to model a participant's probability of selection as being proportional to the participant's personal network size. A relatively new RDS analysis estimator (the successive sampling estimator) assumes sampling *without* replacement and instead models each participant's probability of selection as a function of the participant's personal network size relative to all the personal network sizes remaining in the population after people already in the sample have been removed *(20)*. If the sampling fraction (the target sample size divided by the number of people in the target population) is less than 10% or 20%, the successive sampling estimator may perform similar to the Volz-Heckathorn estimator (RDS II).

Analytical assumptions

Two analytical assumptions of RDS need to be met to make inferences from the sample to the target population:

- respondents can accurately report their personal network size: and
- peer recruitment yields random selection from the recruiter's network.

If these assumptions are not met, the estimators may be biased and thus will not be representative of the target population.

Respondents can accurately report their personal network size

The size of a participant's personal network is the number of personal contacts or peers who are members of the target population. Contacts may include friends, acquaintances, sex partners, coworkers and relatives, as long as they meet the eligibility criteria. To determine a participant's personal network size, survey staff ask a series of questions about the number of eligible people known to the participant. An estimate of personal network size is crucial for calculating sampling weights for RDS, because it determines the probability that a participant will be recruited into the survey. Specific network questions are described further in this chapter under "Additional questions about personal network size".

Peer recruitment yields random selection from the recruiter's network

The assumption of random selection implies that every person in the participant's network has an equal probability of being recruited. Survey staff should not give participants direction about whom to recruit except to say other members of the target population. Random selection may be influenced by factors such as how often peers see each other, how far they live or work from each other, or how well they know each other. Participants may recruit disproportionately from their own subgroup (e.g. they may be more likely to recruit people from their own age group). The assumption of random selection would be violated if one subgroup was more likely to receive or accept coupons than another. However, if the variable associated with biased recruitment (e.g. age group) is not correlated with any of the primary outcome variables (e.g. HIV status), the RDS estimates for those outcomes may not be biased.

The random recruitment assumption is plausible only if members of the survey population have reasonably easy and comfortable access to the survey site, an appropriate time frame is used for the network-size question, and appropriate compensation is provided. However, nonrandom recruitment, if it occurs, will not necessarily bias the RDS estimate if recruitment is not correlated with a variable important for estimation; for example, with the survey's main outcome or degree *(21)*.

Preparing for RDS survey implementation

For an RDS survey to gather a probability-based sample, it must have a way to track recruitment (usually through the use of numbered coupons) and the personal network size of each participant.

Most participants, with the exception of those at the very end of the survey, are given uniquely numbered coupons that they can use to recruit peers. The coupon serves as an invitation to join the survey. The coupon numbers

allow investigators to map the recruitment process and document the chain of recruits stemming from each participant.

A participant's personal network size, or "degree", approximates how many peers the participant can choose to recruit from. Investigators estimate this number by asking each participant how many peers they know who also know them; the investigators can then use this number to adjust for bias and, together with the recruitment information, to estimate the probability of an individual being selected into the survey. All estimates adjusting for RDS assume that these relationships are "reciprocal"; that is, the participants know the peers and the peers know the participants.

With these two measures, investigators can make the necessary analytical adjustments to yield a probability-based sample. If these measures were not factored into the analysis, an RDS sample would still be biased, and considered a convenience sample.

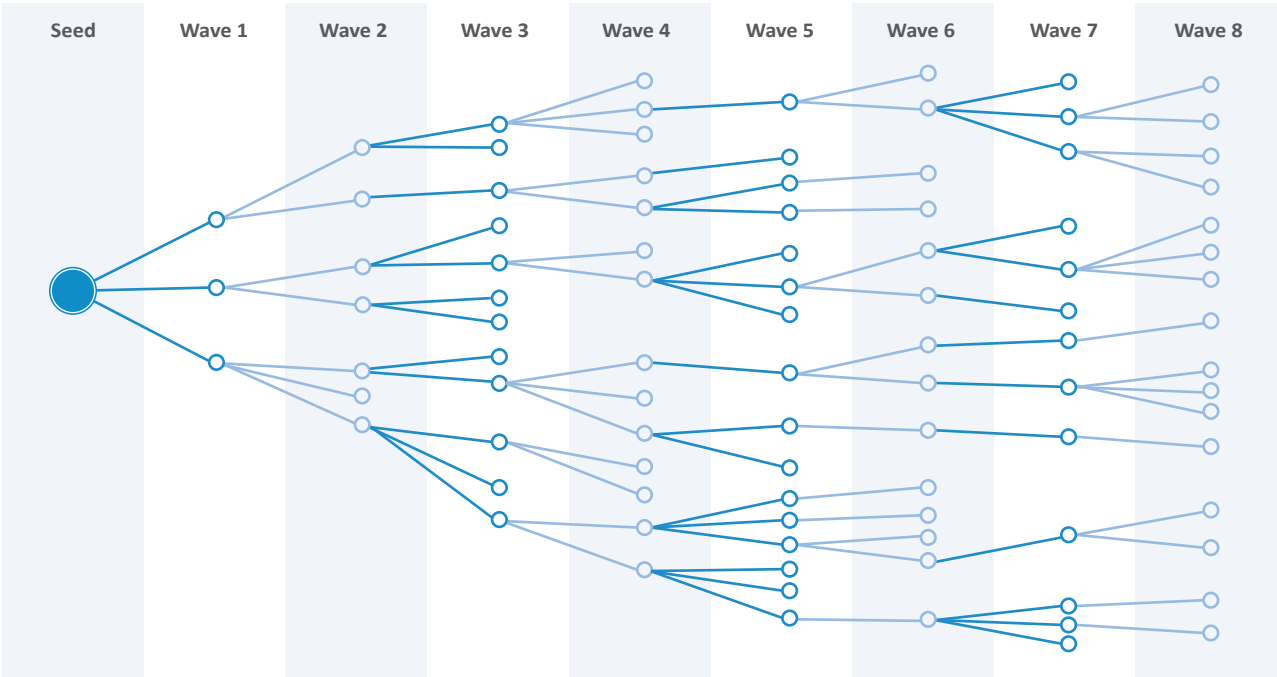
It is not essential to offer respondents compensation for participating in an RDS survey, but it is strongly recommended because participants use their time, effort and money to recruit peers. RDS uses a dual compensation system to encourage participation and recruitment. Primary compensation is given at the end of the first visit. Secondary compensation for the successful recruitment of peers is given during the second visit.

Steps in recruitment

The steps in RDS recruitment are as follows:

1. investigators select and enrol a few individuals from the target population to serve as seeds. Investigators administer questionnaires and biological tests to each seed, and give them a predetermined number of coupons with which to recruit their peers.
2. the seeds use the coupons to invite their peers to join the survey (further information on how many coupons to give each participant is provided below).
3. individuals who received a coupon from a seed report
4. individuals recruited by a seed who then participate in the survey constitute the first wave. In turn, their recruits who then participate in the survey form the second wave (Figure A-9.7).
5. the recruiting process continues until the survey achieves the calculated sample size or higher until the sample reaches equilibrium or converges for the main outcome and salient variables.

Figure A-9.7 Recruitment chain of eight waves generated from one seed



Ideally, the RDS recruitment process generates long recruitment chains, which are needed to reach equilibrium or convergence *(22)*. When equilibrium is reached for a given variable, the sample's values for that variable are stable and no longer influenced by the characteristics of the seeds, which were chosen by investigators, whereas when convergence is reached, the estimators for that variable are stable and are no longer influenced by the characteristics of the seeds. Although it is important to reach a stable composition with respect to key variables, recruitment should continue until the predetermined sample size has been reached.

Not all seeds may be successful in recruiting peers or generating long recruitment chains. It is often the case that only one or two seeds achieve recruitment chains long enough to attain equilibrium and those individuals account for most of the achieved sample size. However, this is not a problem as long as the social network fulfills the assumption of one complete social network (e.g. in Figure A-9.8, that people in the blue chain could have recruited people in the green chain). Figure A-9.8 shows an example of an RDS sample where each successive node and arrow from a seed represents one wave of recruitment. One seed (the grey seed in the figure) produced only one wave of recruitment whereas another (the green seed) produced 13 waves. Because these are all part of one complete network component, the maximum number of waves for any one chain represents the maximum for the entire sample.

Preparations for implementation

Formative assessment results play a key role in informing investigators about how to prepare for implementation.

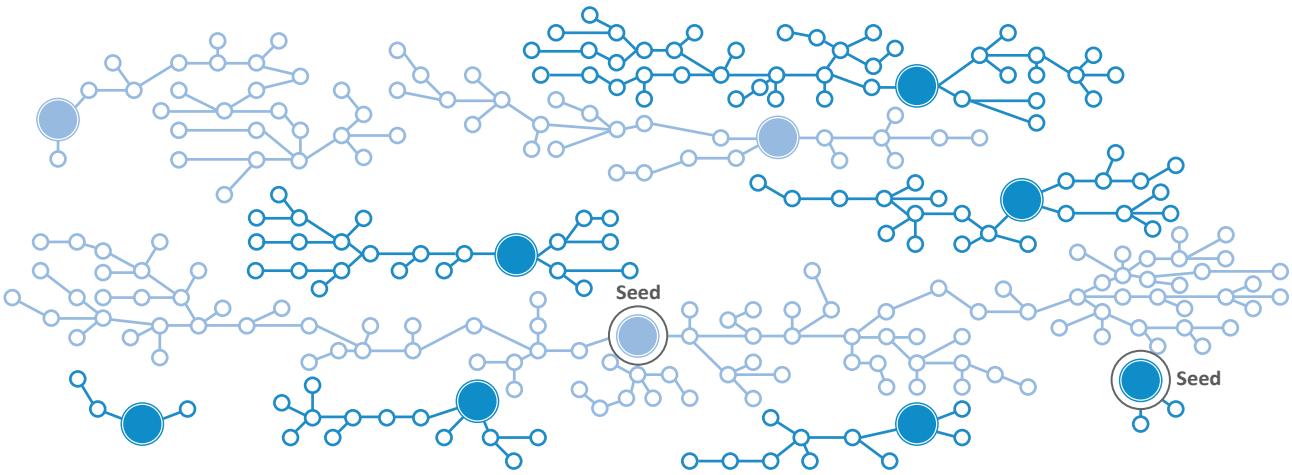
Coupon design and distribution

Coupon design. Coupons are unique survey invitations that participants give to their peers. The coupons are usually paper and can be passed between people. In web-based RDS surveys *(23)* they may also be electronic (Appendix I-21 provides more information about unique participant codes). Regardless of the format, coupons should provide key information about the survey to potential participants, including:

- coupon start and expiry date;
- coupon ID number;
- survey name;
- survey purpose – for surveys in which it is necessary to protect the safety and confidentiality of the participants, this may need to be generic (e.g. "men's health survey" instead of "survey of men who have sex with men");
- survey location, phone or email address;
- hours and days of operation; and
- other important information regarding the survey (e.g. amount of compensation).

Start and expiry dates. Investigators may choose to indicate the start and expiry dates on the coupon. These

Figure A-9.8 Recruitment by 10 seeds



dates inform recruits of when they can visit the survey site. The duration of coupon validity is defined by the investigators and can vary. A future start date (e.g. 1–3 days after a participant's enrolment) may be used to prevent coupon trading and an underground economy (i.e. where peers are waiting outside the survey site and participants immediately recruit them to redeem their secondary compensation).

An expiry date (e.g. 2 weeks after a participant's enrolment) may motivate peers to come to the survey office sooner and may better ensure that a peer's enrolment occurs before the recruiter's second visit during which the recruiter receives compensation for peer recruitment. An expiry date also allows staff to calculate how many valid and expired coupons still circulate "on the streets". Investigators may decide to allow enrolment of persons presenting with expired coupons as long as they meet the eligibility criteria.

Coupon numbering. Each coupon must include a unique ID and be linked to the recruiter. This allows investigators to determine from which recruiter the coupon was received. Coupons should be numbered serially (e.g. 101, 102, 103, 104 and so on). Recruitment should be tracked using a coupon-tracking system such as Respondent-Driven Sampling Coupon Manager (RDSCM). To avoid errors and duplicate coupons, coupons should be numbered and printed before implementation.

Coupon number as survey ID. Ideally in an RDS survey, the coupon code serves as the survey ID, which links interview, biomarker and other data collected from the participant. The survey ID should also be linked to any laboratory code (as applicable) and a unique participant code (UPC) created to confirm participant identity at the second visit (described below).

Figure A-9.9 shows an example of a coupon design.

Figure A-9.9 Coupon design example

Front

MEN'S HEALTH SURVEY

COUPON NUMBER:

Address: 123 Main St. (across the street from Store A)

Open Monday through Friday 10 a.m. – 7 p.m.

Call for more information: XXX-XXX-XXXX

Valid between ____ /____ /____ and ____ /____ /____

Back

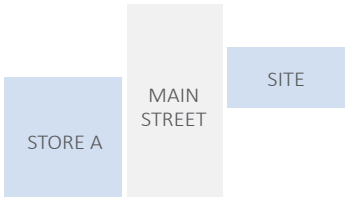
Bring this coupon to join!

If you are eligible, you can:

- Check your health
- Receive free treatment
- Receive compensation for your time

Coupon is not transferrable or valid after expiry date.

Map



Number of coupons. In most RDS surveys, participants are given up to three coupons *(24)*. If participants are given too many coupons, recruitment chains may be wide rather than long and the sample size may be reached before equilibrium or convergence, the point at which the sample is independent from the seeds *(24)*. Therefore, giving more than three coupons per participant should be considered a last option. If recruitment is poor, first consider speaking with participants and others in the community to determine reasons for poor recruitment, undertaking outreach to promote the survey among the target population, improving recruiter training or adding more seeds. The number of coupons distributed should be reduced only when recruitment is robust. If participants are given too few coupons, the recruitment chains may die out.

Survey site selection

RDS surveys are usually implemented at one or more fixed locations, or survey sites. Survey investigators should choose a survey site that is easy to get to via public transportation, is secure, offers privacy for participants, is large enough for all survey procedures and is acceptable to the target population. Formative work can help determine how to improve confidentiality and safety of the site. Participants may feel more comfortable if the site is in a residential area; in other contexts, they may prefer a commercial area or a clinic where other people in the area may assume they are patients.

Accessibility is a matter of location as well as the days and times the survey site is open. If some potential participants have jobs during the day, a survey site that is only open during normal business hours is likely to exclude them from the sample. Some surveys allocate specific days for first visits and second visits. Survey investigators should decide whether setting up appointments or having drop-in times, or some combination of the two, is most appropriate for the target population.

If the recruitment area is large and participants need to travel far to reach any given location, investigators should consider having more than one survey site, and consider the following if selecting multiple survey sites:

- participants may try to participate more than once if the survey operates in multiple sites. In this case, investigators may consider having one survey team that alternates between the sites so that only one site is open per day. Staff may then recognize participants who try to enrol again.
- coupon management requires greater attention to avoid giving out duplicate coupon numbers. One possibility is for each site to provide unique identifiers in the coupon number itself. For example, all "Site A" coupons could begin with the letter A (e.g. A101, A102, A103 and so on) and "Site B" coupons begin with the letter B (e.g. B101, B102, B103 and so on).
- cross recruitment should be monitored to confirm that some recruiters who participated in the survey at "Site A" recruit peers who participate at "Site B".

Mobile sites. Investigators may consider using a mobile site if the survey population is geographically dispersed, or is in an area with a poor or limited public transportation system, or might not visit a fixed site because of security concerns. Mobile sites should have the same characteristics as a fixed site; that is, they should be safe and comfortable, and have a dedicated area for interviewing and specimen collection. In general, a mobile site should be inconspicuous and should not attract attention.

Investigators can operate the mobile site in several locations with assigned hours. Alternatively, they can schedule appointments to meet participants at a convenient public location. The coupons should provide clear instructions about how to find the mobile site. Everyone with a coupon must be given equal opportunity to participate in the survey. Preference must not be given to individuals who live in areas that are more convenient to travel to or visit. All individuals who call to make an appointment should be offered participation if eligible.

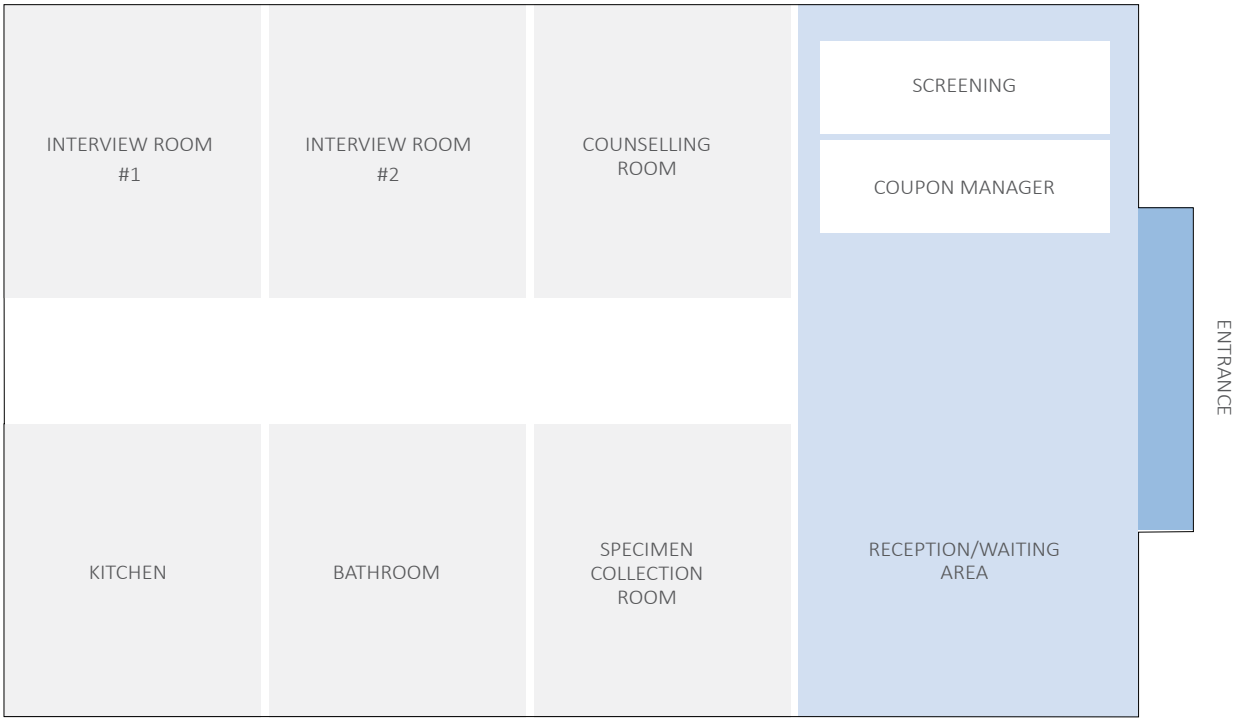
Site layout and staff

Figure A-9.10 shows an example of the ideal layout for an RDS survey site. A survey site should have a reception area where individuals are greeted and can wait between survey procedures. Initial screening for a valid coupon may occur in the same area. Another room or separate space is needed for the coupon manager to screen individuals for eligibility and obtain informed consent. Later, the coupon manager (i.e. an RDS survey staff member responsible for the coupon management system that tracks receipt and distribution of coupons) will explain the recruitment process and give coupons to the participant.

An ideal site has more than one interview room, and has space for a counsellor to conduct pre-test and post-test counselling, for specimen collection and for running any tests. A bathroom should be available for participants as a convenience and is also needed for the collection of urine specimens. A kitchen is useful for preparing refreshments for participants waiting in the reception area. It can also serve as a break room for survey staff. Depending on the biomarkers included in the survey, space for a small laboratory may be needed as well.

Details of what occurs at the first and second visit for each participant are described in Section B-1.4. For more information on survey sites *(25)* and details regarding staffing positions required for an RDS survey, see the information on staff selection in Chapter A-14.

Figure A-9.10 Layout of an RDS site



Seed selection

Seeds are target population members who have been invited to join the survey by survey investigators, to begin the RDS recruitment process. Factors that should be considered when selecting seeds include the number of seeds, the diversity in characteristics of seeds and the likelihood that the seed will recruit others.

Number of seeds. If investigators select too many seeds, the survey site may be overwhelmed with more participants than it can accommodate, and the target sample size may be reached before reaching equilibrium. If investigators select too few seeds, it may take more time to reach the desired sample size. It could also result in some social networks being underrepresented in the sample; for example, people from a certain part of a city. The number of seeds selected may be influenced by the following factors:

- calculated sample size – Achieving a larger sample size may require more seeds.
- compensation – The higher the level of compensation offered, the easier it may be to motivate participation and recruitment; however, this could also lead to ineligible people trying to enrol.
- degree of stigma or hiddenness of the target population – Populations that are subject to more stigma may be more difficult to recruit, in which case investigators may select more seeds if recruitment is not as effective.
- number of connections among survey population members – If the population has many connections (large social networks) this is a good indication that recruitment will be easily sustained. In this situation, fewer seeds may be needed. Seeds should roughly represent the diversity of the target population, including characteristics such as geography, key outcome measure (i.e. HIV positive and negative), socioeconomic status, age and gender. Although diversity is not required, the more representative the seeds are of the target population, the faster the sample may converge towards equilibrium *(10, 26)*. It is also ideal to select seeds that represent a diversity of socially salient variables, which may affect the way each seed recruits. For example, sexual identification may be a factor that influences the way MSM would recruit other MSM and should be considered in seed selection by ensuring that, for example, gay and non-gay identified men are seeds.

Ideal seeds have a large social network and, more importantly, are well connected to their peers, trusted and well liked, and communicate well orally to promote survey participation. Seeds should be well informed about the survey and be enthusiastic about recruiting others to participate. Groups or organizations that work closely with the target population are well placed to help identify seeds. Ideal seeds will be able to recruit across subpopulations to ensure that no bottlenecks occur (i.e. that the sample comprises one complete network component) *(25)*.

Compensation
Although it is not required to offer compensation for participating in an RDS survey, it is strongly recommended. The amount and type of compensation should be based on findings from the formative assessment, and the protocol should include the method for determining the compensation amount.

Investigators should choose an amount of compensation that is sufficient to compensate individuals for their time and effort, but not so much that they participate purely for the compensation, that coupons are bought and sold within the target population, or that people attempt to participate more than once in the survey. Compensation for RDS is best looked at as a remuneration that shows respect for participants' time and effort in recruitment, and their judgement to use remuneration for personal needs *(27)*. A key component in RDS is peer pressure – a mutual like and respect between peers is a strong incentive for participating.

Additional questions about personal network size
Personal network size is the number of people a participant knows who are likely to be eligible for survey participation. It is crucial to assess the size of each participant's personal network in order to weight the data. This information is usually collected through a series of questions that go from broad (e.g. "How many people do you know and they know you who have injected drugs in the past 6 months?") to specific (e.g. "How many [PWID in this city who are at least 18 years old] have you seen in the past 2 weeks?") *(24)*. The data from the last question is used as the personal network size. Because these questions produce the weights used for data analysis, they should be asked face-to-face, even in surveys that otherwise use audio computer-assisted self-interview (ACASI). This makes it easier for participants to ask clarifying questions.

Examples of network size questions for an RDS survey are:

- how many people do you know and they know you who have sold sex in the past 6 months?
- of these, how many live in this city?
- of these, how many of them are aged 15 years or over?
- of these, how many have you seen in the past 14 days?
- of these, how many would you consider giving a coupon to?

Unique participant code
Eligible participants should be assigned a UPC (see Appendix I-21). This code is used at the second visit to verify that the person is the same person who participated in the first visit and was given a specific survey ID. Once a participant's identity has been confirmed, survey staff can compensate the person for recruitment efforts, provide any outstanding test results, and conduct the second visit interview about recruitment efforts.

Investigators can create the UPC by using the answers to a series of questions that only the participant knows the answer to. As such, only the participant will be able to recreate the resulting code (i.e. one that cannot be used to identify the participant). The following pieces of information can be used to create the UPC:

- the first two letters of the participant's first name;
- the first two letters of the participant's mother's name;
- the participant's age in years at the time of the initial interview; and
- the first two letters of mother's place of birth.

In this example, a person named JOHN DOE – whose mother's name is MARY, who is aged 18 years and whose mother was born in RDSVILLE – would have the following UPC.

JOHN DOE MARY 18 RDSVILLE

UPC = JOMA18RD

The code can be created during the eligibility screening and verified by the receptionist or screener to prevent duplicate participation. Alternatively, some surveys use fingerprint scanners that create a unique alphanumeric code without storing the actual fingerprint image, thus retaining anonymity. This code cannot be used to reconstruct the fingerprint. Upon re-scanning the same finger at the second visit, the same code will be generated and the participant's record can be retrieved.

There are guidelines for the reporting of RDS data *(28)*. Also, for guidance on analysis of RDS, WHO has produced a guidance on RDS analysis *(29)*. The University of California in San Francisco has an operations manual for RDS studies, as well as other resources for RDS *(30)*.

A-9.8 References

1 McCreesh N, Frost S, Seeley J, Katongole J, Tarsh MN, Ndunguse R et al. Evaluation of respondent-driven sampling. Epidemiology. 2012;23(1):138.

2 Foreman EK. Survey sampling principles. New York, Marcel Dekker. 1991.

3 Magnani R, Saidel T, Rehle T. Sampling strategies for monitoring HIV risk behaviors. In: Rehle T, Saidel T, Mills S & Magnani R (eds.), Evaluating programs for HIV/AIDS prevention and care in developing countries: a handbook for program managers and decision makers, Arlington, VA, Family Health International. 2001.

4 Munro HL, Pradeep BS, Jayachandran AA, Lowndes CM, Mahapatra B, Ramesh BM et al. Prevalence and determinants of HIV and sexually transmitted infections in a general population-based sample in Mysore district, Karnataka state, southern India. AIDS. 2008;22:S117–S125.

5 Som RK. Practical sampling techniques. Boca Raton, FL, CRC Press. 1996.

6 UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. Guidelines for second generation HIV surveillance: an update: know your epidemic. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/ World Health Organization (WHO); 2013 (http://apps.who.int/iris/bitstream/10665/85511/1/9789241505826_eng.pdf, accessed 3 August 2016).

7 Patton M. Qualitative evaluation and research methods. Beverly Hills, Sage publications. 1990.

8 Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS. 2005;19 Suppl 2:S67–72 (http://www.ncbi.nlm.nih.gov/pubmed/15930843, accessed 30 August 2016).

9 CDC. National HIV behavioral surveillance system: men who have sex with men – Round 4: operations manual. Atlanta, GA: Centers for Disease Control and Prevention (CDC); 2014 (http://www.cdc.gov/hiv/pdf/statistics/systems/nhbs/nhbs-msm4-operations-manual--version-date-05-16-2014.pdf, accessed 4 August 2016).

10 Heckathorn DD. Respondent-driven sampling:a new approach to the survey of hidden populations. Soc Probl. 1997;44(2):174–199.

11 Heckathorn DD. Snowball versus respondent-driven sampling. Sociol Methodol. 2011;41(1):355–366.

12 Van Meter KM. Methodological and design issues: techniques for assessing the representatives of snowball samples. In: Lambert EY (ed), The collection and interpretation of data from hidden populations. NIDA Research Monograph 98, Rockville, MD, National Institute on Drug Abuse. 1990:31–43 (https://archives.drugabuse.gov/pdf/monographs/98.pdf, accessed 8 August 2016).

13 Rutherford GW, Anglemyer A, Bagenda D, Muyonga M, Lindan CP, Barker JL et al. University students and the risk of HIV and other sexually transmitted infections in Uganda: the Crane survey. Int J Adolesc Med Health. 2014;26(2):209–215.

14 Montealegre JR, Risser JM, Selwyn BJ, McCurdy SA, Sabin K. Prevalence of HIV risk behaviors among undocumented Central American immigrant women in Houston, Texas. AIDS Behav. 2012;16(6):1641–1648.

15 Townsend L, Johnston LG, Flisher AJ, Mathews C, Zembe Y. Effectiveness of respondent-driven sampling to recruit high-risk heterosexual men who have multiple female sexual partners: differences in HIV prevalence and sexual risk behaviours measured at two time points. AIDS Behav. 2010;14(6):1330–1339 (http://www.ncbi.nlm.nih.gov/pubmed/20625926, accessed 8 August 2016).

16 Townsend L, Zembe Y, Mathews C, Mason-Jones AJ. Estimating HIV prevalence and HIV-related risk behaviors among heterosexual women who have multiple sex partners using respondent-driven sampling in a high-risk community in South Africa. J Acquir Immune Defic Syndr. 2013;62(4):457–464.

17 Heckathorn DD. Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. Sociol Methodol. 2007:151–207.

18 Goel S, Salganik MJ. Assessing respondent-driven sampling. Proc Natl Acad Sci U S A. 2010;107(15):6743–6747 (http://www.ncbi.nlm.nih.gov/pubmed/20351258, accessed.

19 Gile KJ, Handcock MS. Respondent-driven sampling: an assessment of current methodology. Sociol Methodol. 2010;40(1):285–327 (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3437336&tool=pmcentrez&rendertype=abstract, accessed 3 August 2016).

20 Gile KJ. Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. J Am Stat Assoc. 2010;106(493):135–146 (https://arxiv.org/pdf/1006.4837.pdf, accessed 8 August 2016).

21 Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. Journal of Official Statistics. 2008;24(1):79.

22 Gile KJ, Johnston LG, Salganik MJ. Diagnostics for respondent-driven sampling. J R Stat Soc Ser A Stat Soc. 2015;178(1):241–269 (http://www.ncbi.nlm.nih.gov/pubmed/27226702, accessed 8 August 2016).

23 Wejnert C, Heckathorn DD. Web-based network sampling: efficiency and efficacy of respondent-driven sampling for online research. Sociol Methods Res. 2008:105–134 (http://www.respondentdrivensampling.org/reports/web_rds1.pdf, accessed 8 August 2016).

24 Johnston LG, Malekinejad M, Kendall C, Iuppa IM, Rutherford GW. Implementation challenges to using respondent-driven sampling methodology for HIV biological and behavioral surveillance: field experiences in international settings. AIDS Behav. 2008;12(4):S131–S141 (http://www.ncbi.nlm.nih.gov/pubmed/18535901, accessed 3 August 2016).

25 WHO Regional Office for the Eastern Mediterranean. Introduction to HIV/AIDS and sexually transmitted infection surveillance: Module 4: Supplement. A guide to using RDS Analyst and NetDraw. World Health Organization (WHO) Regional Office for the Eastern Mediterranean/Joint United Nations Programme on HIV/AIDS (UNAIDS); 2013 (http://applications.emro.who.int/dsaf/EMRPUB_2014_EN_1686.pdf?ua=1&ua=1, accessed 8 August 2016).

26 Heckathorn DD. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. Soc Probl. 2002;49(1):11–34.

27 Semaan S, Heckathorn DD, Des Jarlais DC, Garfein RS. Ethical considerations in surveys employing respondent-driven sampling. Am J Public Health. 2010;100(4):582.

28 White RG, Hakim AJ, Salganik MJ, Spiller MW, Johnston LG, Kerr L et al. Strengthening the reporting of observational studies in epidemiology for respondent-driven sampling studies: "STROBE-RDS" statement. J Clin Epidemiol. 2015;68(12):1463–1471.

29 WHO Regional Office for the Eastern Mediterranean. Introduction to HIV/AIDS and sexually transmitted infection surveillance: Module 4: Introduction to respondent-driven sampling [WHO-EM/STD/134]. Eaton, L. A.: World Health Organization (WHO) Regional Office for the Eastern Mediterranean; 2013 (http://applications.emro.who.int/dsaf/EMRPUB_2013_EN_1539.pdf, accessed 8 August 2016).

30 Global Health Sciences. Toolbox for conducting integrated HIV bio-behavioral surveillance (IBBS) in key populations. San Francisco: University of California; 2016 (http://globalhealthsciences.ucsf.edu/pphg/gsi/epidemiologic-surveillance/ibbs-toolbox, accessed 30 August 2016).

Section A

10. Sample size requirements

This chapter describes how to inform and compute sample size requirements for probability-based surveys, including one-time surveys and surveys repeated over time. Although many software packages and online calculators for sample size calculations are available, it is advisable to consult a statistician for such calculations, especially for complex sample designs or the planned use of test statistics in data analysis.

Key terms

Confidence interval:	The estimated range of values which likely includes the true, unknown population parameter.
Design effect (DEFF):	A factor expressing how much larger a sample size for a complex survey design must be compared to simple random sampling. It is the ratio of the observed variance over the (expected) variance if it were based on a simple random sample.
Power:	The probability of not making a type II error (i.e. accepting a null hypothesis that is false).
Precision:	How close an estimate is to other estimates made using the same methodology. The higher the precision, the narrower the confidence interval.
Variance:	A value that indicates the dispersion (scattering) of a variable's values around its average (mean).
Z-score:	Represents the distance between the raw score and the population mean in units of the standard deviation.

In practice, target and actual sample sizes are often influenced both by statistical and resource considerations. Actual sample size achieved may vary depending on the sampling speed, sampling duration or challenges in implementation.

This chapter applies only to surveys using probability samples. For surveys using nonprobability (i.e., convenience) samples, conventional tests of statistical significance do not apply. Instead, practical considerations such as budget and access to the target population will determine the achievable sample size.

A-10.1 Considerations for determining sample size

Biobehavioural survey (BBS) sample size requirements are usually calculated with one of the following goals in mind:

- to estimate the prevalence of a certain characteristic at a single point in time with some specified level of precision (e.g. HIV prevalence among men who have sex with men [MSM] with a defined margin of error); or
- to detect differences between two groups or changes in a certain characteristic over time (e.g. whether condom use at last sex with a client changed among male sex workers (SW) between survey rounds, or whether street-based SW have a lower or higher HIV prevalence than brothel-based SW).

Before calculating a sample size, investigators should consider the types of estimates required, and the variable(s) on which to calculate sample size.

Types of estimates required

The sample size required to detect changes in estimates over time or differences between subgroups is typically larger than that required to estimate the prevalence of a certain characteristic at a single point in time. The purpose of the survey and whether subsequent rounds are expected should be considered when calculating sample size, to ensure sufficiently large sample sizes. Most BBS base their sample size requirements on certain population attributes of interest measured as proportions. Examples include the proportion of participants who test HIV positive, or the proportion of participants who shared injecting equipment the last time they injected drugs. Given the importance of viral load suppression in combating the HIV epidemic, investigators are increasingly basing sample size

calculations around this variable. As a general rule, the smaller the proportion, the bigger the sample size needed (e.g. a "rare event" requires investigators to sample more people in the population to get a precise estimate).

Selecting a variable for sample size calculation

Investigators use a variable that represents the primary objective of the survey to determine the required sample size. If the survey has several primary objectives, investigators determine the sample size needed for each, then select the largest size required. This approach ensures that the target sample size is large enough to satisfy all primary objectives. Viral load suppression in a population is often the most important variable for sample size calculation.

Calculating sample size to estimate prevalence at one point in time

The formula below shows the components of a sample size calculation for a given point estimate. Although we can easily calculate sample sizes with software programs, it is helpful to be familiar with the elements of the formula:

$$n = DEFF*Z^2_{1-\alpha/2}*P*(1-P)/d^2$$

where

n = minimum sample size required DEFF = design effect

$Z_{1-\alpha/2}$ = z-score for the desired confidence level (usually 1.96 for 95% confidence)

P = expected proportion

d = precision

Components of the sample size calculation to estimate prevalence at a point in time

The components of the sample size calculation are expected proportion (P), confidence level ($Z_{1-\alpha/2}$) precision or margin of error, and design effect, as discussed below.

Expected proportion

Estimates (e.g. HIV prevalence) from previous surveys or other sources can inform the value of P. If no previous surveys are available, estimates from nearby areas or countries may be used. Alternatively, investigators may assume (e.g. through a literature review) a certain prevalence ratio for their target population compared to the general population. For example, if no HIV prevalence estimate for male SW is available, investigators may multiply the (known) HIV prevalence among the general male population in the same area by a factor of 2 or more, based on data from other studies.

Confidence level

The confidence level refers to the percentage of all possible samples that can be expected to include the true population parameter. For example, a 95% confidence level implies that 95% of the confidence intervals include the true population parameter.[1] In general, the narrower the confidence interval chosen, the larger the required sample size.

The confidence level is often calculated based on the z-score distribution, which indicates how many standard deviations a value is from the mean. A 95% confidence level (implying an alpha of 0.05, meaning there is a 5% likelihood that the true population parameter lies outside the 95% confidence interval) corresponds to a z-score of 1.96. The z-score for a 90% confidence level is 1.645, whereas that for a 99% confidence level is 2.326.

Precision or margin of error

Because a sample only provides an estimate of the true population parameter, the precision of a sample reflects the reproducibility of the survey's measurement using the same methods. One measure of precision is the standard error, with precision being inversely related to the standard error.[2] The higher the precision, the narrower the confidence interval. For example, for a survey with an assumed prevalence of 50% and a confidence level of 95%, a precision of 0.10 (10%) means that 95% of the time the survey (with a given sample size) will yield an estimate falling within 40–60% (the confidence interval); that is, 50% ±10%. When calculating the sample size, the precision needs to be used as an absolute value (i.e. 50% ±10% in the above example) and not as a relative value (i.e. 20% in the above example).

Design effect

A design effect is included in the sample size formula when a probability survey other than a simple random sampling survey is used. It quantifies how much larger the sample size needs to be to provide the same precision as a simple random sample, and is a multiplier used to inflate the sample size by a certain value. A design effect of 2, for example, means that a sample size of 400 needs to be doubled to 800 in order to achieve the same precision as a simple random sample.

A design effect inflates the necessary sample size compared to a simple random sample because the "variance" in such samples is larger than in simple random samples. Variance indicates how far a set of numbers (values) is spread out in the population. The design effect reflects the ratio of the actual (observed) variance to the expected variance if a simple random sample were used. Probability-based nonrandom

[1] Adapted from: http://stattrek.com/statistics/dictionary.aspx?definition=confidence_level

[2] Adapted from: http://stattrek.com/m/statistics/dictionary.aspx?definition=precision

sampling designs (e.g. cluster, time-location or RDS) produce less varied samples, leading to wider confidence limits. In other words, less varied sampling designs provide less precision than a simple random sample; therefore, such nonrandom sampling designs lead to smaller "effective" sample sizes.

Importantly, the design effect varies by the variable of interest (participant characteristic); for example, the design effect to measure HIV prevalence with a certain precision may be different from the design effect to measure ever having been tested for HIV.

Investigators need to be careful not to confuse the design effect with the design factor (DEFT), which is the square root of the design effect. Sometimes, survey reports publish DEFT rather than the design effect. In these cases, the DEFT value (e.g. 1.5) needs to be squared in order to derive the design effect (in this example, 1.5 × 1.5 = 2.25) and so obtain the factor by which the effective sample size needs to be inflated.

How to calculate the design effect

The design effect cannot be predicted with accuracy before a survey, and therefore needs to be estimated based on recommendations or data from previous surveys. After a survey, an estimate of the design effect can be calculated by dividing the observed variance by the variance based on a simple random sample. Calculating the design effect after the survey informs future surveys and gives a sense of whether the assumed pre-survey design effect value was accurate.

Design effect in respondent-driven sampling (RDS) surveys

In RDS surveys, design effect values not only vary substantially across surveys and within surveys by indicator, they may also be high. In the past, a design effect of 2 was commonly anticipated for RDS surveys. However, recent theoretical work suggests that a larger design effect should be expected (e.g. 3 or 4) (*1, 2*). Although such a large design effect can dramatically reduce the effective sample size (i.e. the sample size corresponding to that of a simple random sample), investigators may have few alternatives available other than accepting a lower precision, striving for the largest sample size possible or opting for a different sampling method. The two sources of uncertainty in RDS estimates are bottlenecks in recruitment (i.e. recruitment across distinct and poorly connected networks leading to segmentation of the target population) and nonrandom recruitment within recruiters' networks.

Investigators should refer to the current literature for theoretical examinations of expected design effect values as well as observed design effect values in comparable RDS surveys.

Design effect for cluster-based surveys

In conventional cluster sampling (CCS) or time-location sampling (TLS) a design effect around 2 should be used. Results within clusters are typically expected to be correlated, which requires that the sample size be increased when compared to that needed under simple random sampling.

Example sample size calculation for a single proportion

In an RDS survey designed to estimate HIV prevalence among people who inject drugs (PWID), the expected HIV prevalence (P, informed by a previous survey) is 20% (P = 0.2), if the design effect is set at 3, the confidence level at 95% ($Z_{1-\alpha/2}$ = 1.96) and precision at 5% (d = 0.05):

Design effect = 3
$Z_{1-\alpha/2} = 1.96$; $(1.96)^2 = 3.84$
P = 0.20
d = 0.05; $(0.05)^2 = 0.0025$

Using these data, the final sample size is 738 (rounded up):

$3 \times (3.84 \times 0.2 \times (1–0.2))/0.0025 = 738$

Other elements to inform sample size calculation

Sometimes, investigators are interested in a particular subgroup of the population – such as users of a particular drug, SW who only work outside brothels or TG – which may make up only part of the survey sample. If a certain precision is required for an estimate in a subgroup, then the total target sample size needs to be larger. Investigators need to estimate the relative size of such a subgroup (e.g. 50%), calculate the sample size for the intended precision, and then "inflate" the total sample size by that factor (e.g. a subgroup comprising 50% of the target population would imply a doubling of the sample size).

Limitations of the presented method for sample size calculation are discussed below.

Possible underestimation of required sample size

The presented method is somewhat non-conservative (i.e. may suggest lower sample sizes than warranted), especially where assumed prevalence values (proportions) are low. In such situations, other approaches (e.g. Wilson score or asymptotic bounds with continuity correction) are preferable and would typically give larger sample sizes. As mentioned earlier, a statistician should be consulted regarding the choice and appropriateness of the sample size formula.

Missing data and the effect on sample size

Even when surveys meet their target sample size, the resulting datasets may be incomplete due to accidental

data loss (missing data), erroneous values, refusal to provide an answer (nonresponse), or loss of specimens or refusal to provide a biological specimen. If the possibility of incomplete or erroneous data or missing specimens cannot be excluded during survey planning, a corresponding increase in sample size is recommended to offset their effect. To do this, the original sample size is divided by (1-x/1), where x represents the proportional extent of nonresponse. For example, if investigators fear that 10% (= 0.1) of the target 1000 specimens will not yield biomarker data due to factors such as loss, then the corrected sample size is derived through 1000/((1-0.1)/1), yielding 1111.

Lack of finite population correction

The formula shown above for sample size calculation assumes that the sample size is much smaller than the population size. This means the total population is considered "infinite"; that is, much larger than the sample size. If this is not the case, a finite population correction should be used.

The final population correction (fpc) is computed as $f\,pc = \sqrt{((N - n)/(N - 1))}$, **where *N*** denotes the total population size, and n denotes the sample size. The computed fpc is then used to refine the expected standard error, which determines the confidence level. For example, if the total population is 10, 000 and the sample size is 600 the finite population correction would be:

$$\sqrt{((10,000-600)/(10\,000-1)} = 0.9695 \approx 0.97$$

To use a finite population correction, the total population size must be known, but often it is not. Hence, total population size is often not considered when determining sample size. In all cases, investigators should be confident that the computed sample size is achievable given the total population size, meaning that the total population size should be substantially larger than the sample size.

Once data collection has started, investigators should examine the incoming data early on to see whether the assumptions are likely to be met; this may allow investigators to redo the sample size calculations and alter the target sample size in time.

Calculating sample size to detect changes over time

An alternative framework for calculating sample size is based on looking for a change in prevalence over two survey periods rather than a desired precision in the estimate for one survey.

Detecting changes in proportions

For a variable of interest, P_1 and P_2 are the proportions expected at Time 1 and Time 2. For example, investigators may wish to detect a change in HIV prevalence among SW from 30% (P_1) to 20% (P_2) or less some time later.

Selecting P_1 and P_2

P_1 should be determined as described above (see "*Components of the sample size calculation to estimate prevalence at a point in time*"). P_2 is set at the target proportion (20% or 0.2 in this example). In practice, it is usually set at the smallest meaningful change investigators expect might have occurred. For example, a decline in HIV prevalence from 30% to 20% may be considered meaningful, but a decrease from 30% to 28% may not. The larger the difference between the two proportions, the smaller the sample size needed to determine a statistically significant difference. Investigators should consider P_1 and P_2 carefully. Programmatically, it is often easier to achieve a drop in HIV prevalence from 30% to 20% (a drop of one third) than from 15% to 5% (a drop of two thirds), even although the absolute difference is the same.

Consider also the time interval and type of outcome measurements. It is difficult or impossible to detect substantial changes in HIV prevalence in a short period of time (e.g. a drop in HIV prevalence from 30% to 20% over 2 years is unlikely to occur), whereas changes in behaviours are much likely to be detected in the short term (e.g. uptake of HIV testing or condom use).

In addition, the closer P_1 and P_2 are to 50%, the larger the sample size needed to meet a desired precision.

The sample size required for each survey for the measurement of change between two survey rounds for a given indicator is a function of four factors:

- initial prevalence of the characteristic;
- size of the difference to be detected;
- level of significance, or the degree of certainty that a change of that magnitude would not have occurred by chance; and
- power, or the degree of certainty that a change of that magnitude can be observed if it actually does occur.

The formula for calculating the sample size for each survey round (n) is given by:

$$n = DEFF \times (Z_{1-\alpha/2} \times \sqrt[2]{2p(1-p)} + Z_{1-\beta} \times \sqrt[2]{p_1(1-p_1)+p_2(1-p_2)})^2 / (P_2 - P_1)^2$$

where

DEFF = design effect

P_1 = the estimated proportion at the time of the first survey

P_2 = the target proportion at some future date (second survey), so that ($P_2 - P_1$) is the size of the difference that is to be detectable

$P = (p_1 + p_2)/2$

$Z_{1-\alpha/2}$ = the z-score corresponding to desired level of significance, usually 1.96 for 95%, and α is the type I error rate (reflects the probability of the test suggesting that there was a change in the proportions when there was not one).

$Z_{1-\beta}$ = the z-score corresponding to the desired level of power, usually 0.83 for 80%, and β is the type-II error rate (reflects the probability of the test suggesting that there was no change in the proportions when there was one).

As indicated by the z-score, this formula is for a two-tailed test; that is, the required sample size will enable the detection of a difference in P in either direction (down or upwards).

Statistical power

The statistical power score ($Z_{1-\beta}$) corresponds to the power required to detect change over time. Usually, 80% power is selected, meaning there is an 80% chance that a survey will detect a change in a certain proportion over time or between groups, if there actually is a change. The ability to detect a difference will depend on the sample size, the magnitude of the difference and the variance. The more statistical power needed, the larger the resulting sample size.

Examples of sample size calculations to detect a change in proportions

Example 1.
Suppose investigators are planning a survey of female sex workers (FSW). The goal is to measure whether condom use at last sexual intercourse with a client will increase from 20% in the baseline survey to 30% or more in the next survey round. They set the design effect at 2 (DEFF = 2), confidence at 95% ($Z_{1-\alpha/2}$ = 1.96, for a two-sided test[3]) and power at 80% ($Z_{1-\beta}$ = 0.84). Using these parameters, the final sample size is 587:

$$2 \times (1.96 \times \sqrt{(2 \times 0.25(1–0.25))} + 0.84 \times \sqrt{(0.20 \times (1–0.20) + 0.30 \times (1–0.30))})^2 / (0.30–0.20) = 587$$

In this example, investigators would need to conduct two surveys with a sample size of at least 587 participants in each survey.

Example 2.
Suppose investigators want to detect a decrease of 15 (absolute) percentage points in the proportion of transgender male vocational students who had unprotected sex in the past 12 months. A level of significance of 95% and a power of 80% is desired. On the basis of earlier survey data, it is thought that the appropriate "baseline" value on the indicator would be 55%.

Set P_1 = 0.55 and P_2 = 0.40, and use z-score values of $Z_{1-\alpha}$ = 1.645 (95% significance level for a one-sided test) and $Z_{1-\beta}$ = 0.84 (corresponding to 80% power) and obtain:

$$n = 2\,[1.645\,\sqrt{2(0.475)(0.525)} + 0.84\,\sqrt{(0.4)(0.6) + (0.55)(0.45)}]^2/(0.40–0.55)^2$$
$$= 2 \times [(1.1617+0.5865)2/0.0225] = 272$$ vocational students in each survey round.

Note: Sample design assumes a design effect of 2.

Appendix I-22 shows examples of sample size requirements for a range of different scenarios. The same limitations and considerations as shown for one-time surveys above apply.

A-10.2 Additional considerations

Considerations for web-based surveys

Many web-based surveys are considered convenience samples, so considerations such as statistical power and precision may not apply. Web-based surveys attempting a sampling design to yield some form of probability sample (e.g. a web-based RDS, representative of MSM accessing the internet) should follow the methods described in this chapter to calculate sample size *(3, 4)*.

Time and budget

Sampling speed may have a substantial influence on achieving the sample size, because many survey cost elements are time dependent (e.g. staff time, rent, utilities and transport) rather than a function of the total target sample size. Hence, the more participants a survey can enrol each day or week, the more cost effective it generally is and the higher the resulting sample size. For example, an RDS survey that starts recruiting slowly and fails to increase the speed of recruitment will yield a correspondingly smaller total sample size by the time the budget is exhausted than a survey that enrols far more respondents each week.

[3] A two-sided test it is one that can detect a change in either direction (i.e. either an increase or a decrease in this example). A one-sided test is one that is only powered to detect a change in one direction. Unless it is clear that a change in the outcome of interest can only be in one direction, a two-sided test should be used. The second example uses a one-sided test because investigators are only interested in detecting a decrease, rather than a change in either direction. The sample size required for a one-sided test is smaller than for a two-sided test.

Calculating the size of a second BBS

Often the decision to conduct a survey to detect a change compared to a baseline value comes after the first survey has been conducted. In such a situation, investigators should use the first survey's results to inform the second survey, including the baseline value (proportion) and surrounding confidence limits, as well as the design effect. This allows investigators to have more confidence that their second survey's target sample size will be able to detect the anticipated difference.

Sample size calculators

Many statistical software packages include tools for sample size calculations, such as EpiInfo.[4] A simple Excel-based calculator for cluster-based surveys can be found on the WHO website.[5] Various sample size calculators are included in Appendix I-22 and online.[6]

A-10.3 References

1 Johnston LG, Chen Y-H, Silva-Santisteban A, Raymond HF, DiNenno E. An empirical examination of respondent-driven sampling design effects among HIV risk groups from studies conducted around the world. AIDS Behav. 2013;17(6):2202–2210.

2 Wejnert C, Pham H, Krishna N, Le B, DiNenno E. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. AIDS Behav. 2012;16(4):797–806 (http://www.ncbi.nlm.nih.gov/pubmed/22350828, accessed 8 August 2016).

3 Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy. J Stud Alcohol Drugs. 2012;73(5):834–838 (http://www.ncbi.nlm.nih.gov/pubmed/22846248, accessed.

4 Bengtsson L, Lu X, Nguyen QC, Camitz M, Hoang NL, Nguyen TA et al. Implementation of web-based respondent-driven sampling among men who have sex with men in Vietnam. PLoS One. 2012;7(11):e49417 (http://www.ncbi.nlm.nih.gov/pubmed/23152902, accessed 8 August 2016).

[4] http://www.cdc.gov/epiinfo/

[5] http://www.who.int/chp/steps/resources/sampling/en/

[6] http://statpages.org/

11. Population size estimation methods used with surveys

In many countries, the HIV epidemic was first identified among key populations. The key populations covered in these guidelines are often "hard-to-reach" or "hidden" populations because their behaviours are illegal or stigmatized in some settings; hence, the size of these populations is usually unknown. This chapter discusses methods used to estimate population size, because knowing the size of a key population can help investigators to understand the scale of the response needed to ensure the population gets services to prevent HIV infection.

Purpose or objective of population size estimation

Estimating the size of key populations at risk for HIV is important in a number of ways. For example, size estimation data can be used to:

- inform policy and advocacy;
- provide denominator data for indicators related to reaching coverage for key populations;
- provide critical information for models used to estimate and project HIV impact;
- inform HIV response planning, target setting and resource allocation (e.g. funding or budget); and
- inform service delivery and facilitate programme monitoring and evaluation (e.g. programme coverage) *(1, 2)*.

A biobehavioural survey (BBS), in conjunction with other studies, can provide an opportunity to estimate the size of the surveyed population.

The Joint United Nations Programme on HIV/AIDS (UNAIDS) global, regional and national HIV estimates for all but generalized HIV epidemic settings use key population size estimates to estimate the number of new and prevalent infections, the number of people eligible for care and treatment, and the number of HIV-related deaths. Population size estimates have also been used to:

- mobilize political support and commitment for the HIV response;
- direct funding by characterizing the extent of the epidemic;
- plan HIV programmes for key populations nationally, regionally and locally; and
- monitor and evaluate programmes in terms of coverage, quality and effectiveness.

Results of a size estimation study can garner much-needed media attention to demystify often-held notions that the actual number of certain key populations is negligible and hence not worthy of public health action. Population size estimates can be used as denominators for reporting on international monitoring indicators, and for grant applications.

Population size estimation methods currently in use

A number of size estimation methods are available for use. No gold standard exists. All methods have their strengths and limitations, and where possible, multiple estimates should be produced to derive a consensus estimate. Table A-11.1 presents an overview of different size estimation methods. The available methods can broadly be categorized into direct methods (census and enumeration) and indirect methods. Direct methods count members of the population directly, whereas indirect methods use data from different sources to estimate their size. In addition, these methods can be classified through their data source: data collected from at-risk populations, and data collected from the general population.

Table A-11.1 Summary of methods for estimating the size of key populations

Method	Description	Strengths	Assumptions/limitations
Census	Count all members of the population	• Real count, not an estimate or sample • Can produce credible lower limit • Can be used to inform other methods	• At-risk populations are often hidden; misses some or many members of the population • Stigma may prevent members from identifying themselves • Time consuming and expensive
Enumeration	Develop a sampling frame and count all members of the population at the selected locations	• Can produce credible lower limit	• At-risk populations are often hidden; misses some or many members of the population • Stigma may prevent members from identifying themselves
Capture-recapture	Size estimate is based on two independent captures (samples): • capture 1: "tag" and count number tagged • capture 2: "tag" and count number "retagged" (i.e. who had also been "tagged" the first time)	• Relatively straightforward to do with access to population • Does not require much data	Relies on four conditions that are hard to meet: • the two captures must be independent and not correlated • each population member should have equal chance of selection • each member must be correctly identified as "capture" or "recapture" • no major in or out migration
Multiplier method	Apply a multiplier (e.g. number receiving particular service or having membership, or number receiving a unique object distributed before a survey) to survey estimate (proportion of survey sample sharing same characteristic)	• Uses data sources already available • Flexible in terms of sampling methods; first source need not be random, but second source should be representative of population	• The data sources must be independent • The data sources must define population in the same way • Time periods, age range and geographical areas must be aligned • Data collected from existing sources may be inaccurate
General population-based survey	Ask respondents if they engage in the behaviour of interest (e.g. male–male sex, money for sex, inject drugs) Generates proportional estimates (% of the general population who engages in a certain high-risk behaviour)	• National surveys are common and familiar • Easy to implement if a survey is underway • Straightforward to analyse • Sampling is easy to defend scientifically ("gold standard")	• Low precision when the behaviours are rare • Respondents may be reluctant to admit to stigmatized behaviours • Only reaches people residing in households (mobility) • Privacy and confidentiality risk to subjects

Method	Description	Strengths	Assumptions/limitations
Network scale-up method (NSUM)	The general concept is that an individual's social network is representative of the whole population (i.e. one person's group of friends somehow reflects the characteristics of the whole community); thus, members of the general population are asked whether their acquaintances have high-risk behaviours (e.g. buying or selling sex, having anal sex between men or injecting drugs). The average proportion of the respondents' acquaintances who have these behaviours is used to estimate the proportion of the adult population with these behaviours.	• Can generate estimates from general population rather than hard-to-reach populations • Does not require survey respondent to disclose stigmatizing behaviours • Clear formulation and calculations have been developed and tested • It is possible to get confidence intervals	• Average personal network size is difficult to estimate • Some subgroups may not associate with members of the general population • Transmission error: respondents may be unaware someone in their network engages in the behaviour of interest • Reporting bias (i.e. social desirability) may arise • May work well with some groups and in some contexts but not in others
Reverse tracking method	Compares the "observed size" with the "estimated size" for a selected site, calculates an adjustment/correction factor and modifies the total estimated size in the sampling frame accordingly	• Does not require collection of additional information outside of a TLS survey	• Requires an exhaustive list of sites • Would not be able to estimate the hidden part of the population (i.e. that portion that never accesses these sites) • Could result in overestimation if the same individuals are found in many sites
Successive sampling population size estimation	A probability model is created as a function of the observed personal network sizes in the RDS sample, investigators' knowledge about the population size, and the unobserved network sizes	• Does not require collection of additional information outside of the BBS • Can be conducted with multiple BBS sampling methods including TLS and RDS	• Estimate's validity depends heavily on representativeness of the sample • Prior information on the population size may be unavailable or poorly estimated
One-sample capture-recapture method	Requires sampling from each participant's network connections and matching those connections against the other participants in the sample and the list of their respective contacts	• Does not require collection of additional information outside of an RDS survey	Truncated Poisson estimates for sparse data rely on certain assumptions, some of which may be hard to meet: • a closed population (no entry or exit); • behaviorally homogeneous population members; and • a constant encounter rate, with no behavioral response to an encounter that reduces or increases the rate for future encounters

Sources: Abdul-Quader et al 2014, UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance, 2010 (*2, 3*)

For a detailed description of each of these methods, please refer to the recent UNAIDS/WHO guidelines document on size estimation *(2)* and other relevant literature in the references.

A-11.1 Population size estimation based on surveys

Because these guidelines focus on integrated BBS, among the methods listed above, two methods can be integrated within a BBS: the multiplier method and the capture-recapture method (CRC). The remainder of this chapter covers key assumptions and strengths and limitations of these two methods as well as emerging methods, and considerations for selecting a method.

A-11.1.1 Multiplier method

The multiplier method compares two independent sources of data to estimate the total number in a population. The first source is a count or listing from programme data including only the population whose size is being estimated, and the second source is a representative survey of the same population *(1)*.

Service multiplier

- data source 1 – the count: Review the programme interventions and identify specific programme services. From programme data, count the total number (B) of the population (whose size is being estimated) who received that particular service within a specified geography and given time period.
- data source 2 – the multiplier: Estimate the proportion (m) of the population (whose size is being measured) receiving that particular service during the same reference period through a representative survey among the target population within the same geographical region. This is known as "service multiplier".
- multiply the count (B, data source 1) by the inverse of the proportion of the population (1/m) who say they received services over the same period using the following simple mathematical formula:

$$N = B/m$$

where

N = estimate of total population size

B = total number of the population who received a particular service (programme data)

m = proportion of the population receiving a particular service (survey data).

C2 = sample size of the second capture (i.e., the survey)

Programme data can be from sources such as outreach HIV testing services, needle or syringe exchange programmes, or sexually transmitted infection (STI) clinics.

The variance can be computed as:

$$Var(N) = [B*C2*(B-m) \times (C2-m)]/m^3$$

The 95% confidence intervals can be computed as:

$$95\%CI = N \pm 1.96 \times \sqrt{Var(N)}$$

Example of service multiplier method

A representative survey of 800 FSW was conducted in a city. Participants were asked whether they had visited a certain key population health centre in the past six months. The survey found that 11.0% (88/800) of participants had visited the health center in the past six months. The health centre reported having 395 individual FSW patients in the past six months. Thus, the estimated number of FSW in the city is:

B=395
m=88/800=11.0%
N = B/m = 395/0.11 = 3,591

Object multiplier

Another version of the multiplier method involves distribution of a unique object (i.e. one that cannot be bought easily or otherwise obtained) to the target population before the survey. Investigators often discuss a suitable object with representatives of the survey population (key informants) to ensure the object is acceptable and has intrinsic value (utility); that is, the object is useful to keep, but does not have commercial value that may facilitate its sale or trade for other things. Bottle openers, key chains or nail clippers are such objects. Once identified, investigators will often order a set number of custom-made objects. The object is distributed widely in the sampling area and only to those individuals likely to meet the survey's eligibility criteria. Recipients of the object are told that they should not lose it or give it to someone else. When the survey participants are recruited and interviewed, they are asked whether they received the object that was previously distributed. If any survey participants are not able to show the object that was distributed before survey recruitment, they are shown a number of similar objects and asked to identify the correct one. The total number of the survey participants who correctly show or identify the object they had received and the proportion of them recruited to participate are used to estimate the size using the formula presented above. Investigators should aim to distribute at least twice as many objects as the target sample size. Given that this may not always be feasible, investigators should distribute as many objects as possible. The more objects distributed, the more precision can be expected in the estimate. Appendix I-23 includes a calculator to determine the number of unique objects required to obtain a given level of precision.

Event multiplier

Another version of the multiplier method involves conducting an event for the target population before the implementation survey. When the survey participants are recruited and interviewed, they are asked if they had attended the specific event that was held on a certain day and time at a certain location. The total number of survey participants who attended the event and the proportion of them recruited to participate are used to estimate the size using the above formula.

Assumptions

There are four key assumptions to be fulfilled for the multiplier method:

- the data are for unique individuals (i.e. no double counting);
- both data sources are using the same definition for the population;
- data from two sources are available for the same time period; and
- both data sources have the same catchment area (geographical coverage).

Strengths and limitations

Multiplier methods are relatively straightforward to use and perhaps the most widely used of size estimation methods *(1)*. However, they are subject to the quality of the service (event or object) data. Poor service data or programme data that are challenging to match with the survey's eligibility criteria in terms of geographical coverage, risk behaviour and age often present a limitation to the quality of the resulting size estimates. Further, investigators need to make sure that the service data can provide numbers on unique persons reached (i.e. exclude persons who show up in the service data multiple times). The survey questions related to the size estimation need to be very clear (i.e. sensitive and specific). Questions should ask about the specific service that is being used to estimate the size.

Unique object multipliers may be the choice where service data is nonexistent or of poor quality. The greatest difficulty in using multiplier methods is finding data for institutions and populations that correspond with one another in terms of the definition of population, time reference period and, particularly, catchment area. Because of the catchment area issue, multiplier methods based on service use are most commonly used at the local level. The multiplier method based on unique object distribution could be used at a larger level or in an area without good institutional data; however, it relies on access to members of the key populations in which a particular unique object is distributed as widely as possible. Multipliers based on unique events rely on having the survey population aware of the event and attending the event without the fear of stigma and discrimination.

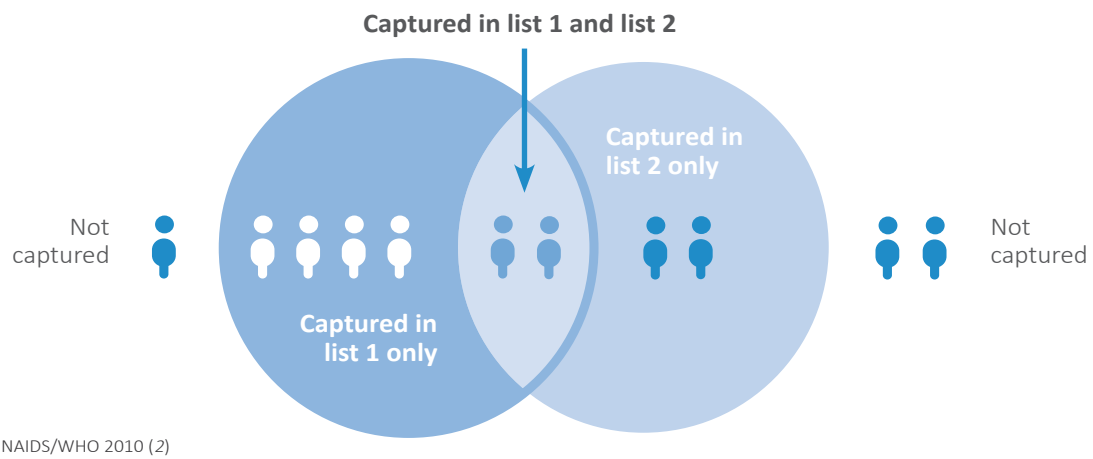
Whatever multiplier is chosen, the key assumptions should be met and both data sources (multiplier and survey) should be representative of the survey population, otherwise, the resulting estimates are subject to bias.

A-11.1.2 Capture-recapture method

CRC originated in biology (counting fish in a pond), and until recently it was used primarily to estimate wildlife populations (where it was known as the Lincoln-Peterson estimator). Other terms sometimes used include "mark and capture" and "capture and release" *(4-10)*.

Figure A-11.1 illustrates CRC.

Figure A-11.1 Capture-recapture method



Source: UNAIDS/WHO 2010 (*2*)

How capture-recapture works

- prepare a list – as exhaustive as possible – of the sites where the population can be found, based on information collected during mapping or sampling frame development exercises of BBS. Visit all the sites on a given day within a certain time period, and tag all members of the population found at the site (give them some unique identifier, such as a card or memorable gift). Keep a count of the persons tagged (C1).
- revisit all the sites 1 or 2 weeks later within the same time period, and tag all members of the population encountered at the second visit. Ask all those encountered (approached) whether they were tagged (received the unique identifier) in the first visit (1–2 weeks earlier) by someone from the team. Keep a count of the persons tagged at the second visit (C2).
- identify and count the persons who were tagged twice, both in the first and second visit (R).

The three key counts mentioned above: total number of persons captured during the first visit (C1), total number of persons captured on the second visit (C2), and total number of persons captured twice (R) are used to estimate the size using the following simple mathematical formula:

$$N = (C1 \times C2)/R$$

where

N = estimate of total population size

C1 = total number of persons captured on the first visit

C2 = total number of persons captured on the second visit

R = total number of persons captured on both the visits.

The variance can be computed as:

$$Var(N)=[C1*C2*(C1-R)*(C2-R)]/R^3$$

The 95% confidence intervals can be computed as:

$$95\%CI=N\pm1.96*\sqrt{Var(N)}$$

Assumptions

There are four key assumptions that need to be fulfilled for applying CRC:

- the population is closed (i.e. no significant in or out migration);
- individuals captured on both occasions can be matched (i.e. no loss or misclassification of marks);
- for each sample, each individual has the same chance of being included (i.e. same catchability); that is, people in the first sample are not more or less likely to be included in the second sample than people who were not included in the first sample; and
- capture in the second sample is independent of capture in the first (i.e. the two samples are independent); that is, either key population members are tagged at all sites, or investigators draw a random sample of sites on each occasion (investigators using a sample of sites should not go to the exact same sample of sites both times).

CRC can also be conducted by sampling a selected number of sites (locations) when the size estimation area is much larger and the total number of sites (locations) are numerous. In this instance, a random sample of sites (locations) are selected for the first capture. For a recapture exercise, another sample of sites (locations) are selected randomly. This helps with drawing independent samples for R and C.

Example of the use of capture-recapture method

The following example summarizes the steps involved in a population size estimation activity conducted to estimate the size of the population of street-based sex workers (SW) in Dhaka, Bangladesh.

Step 1:
Existing data were reviewed and analysed to understand the nature of street-based sex work, including characteristics, estimated size, locations, accessibility and operating hours.

Step 2:
The population was defined; in this case, FSW who negotiated with their clients on the street.

Step 3:
The geographical area was defined, to be included for the estimation.

Step 4:
The defined geographical area was divided into a number of zones and field staff were assigned to each zone.

Step 5:
Ethnographic observation and mapping of each of the zones were conducted to assess the following:

- prevalence of street-based SW within that zone;
- locations or venues within that zone where the street-based SW were likely to congregate and where they could be approached for counting during the estimation exercise;
- when they can be found and counted (day of the week and time of the day, e.g. early evening or late evening); and
- estimated number of SW that congregated at each of the identified locations within the zone.

In addition to conducting ethnographic observation, the field staff conducted brief informal interviews with local

business owners to verify information collected based on observation and mapping, including estimated number of street-based SW within that zone.

Step 6:
SW were recruited as volunteer field staff from known sources (e.g. from existing STI clinics).

Step 7:
Orientation was provided to the SW about the purpose of the project activity.

Step 8:
Field teams were formed for each zone, with each team including at least one project field staff and two to three SW recruited as volunteers.

Step 9:
Field teams were trained on the size estimation method (CRC) including the principles and operations of the method.

Step 10:
Based on information and secondary data collected, it was decided that the first count (first capture) would be conducted on a given day between the times of 7:00 pm and 3:00 am.

Step 11:
About 8000 small cards were printed for distribution during the first count, which was based on existing information on street-based SW in the city. The cards were sequentially numbered from 0001 to 8000. In addition, the cards had information about one of the STI clinics (address, telephone number and services available) that provided STI services to SW. Each field team received a set number of cards, and records were kept. The field team was instructed to keep track of the number of cards distributed and those not distributed and returned to the project office. This was done to determine the total number of cards that were distributed in the first count.

Step 12:
On the day that was selected for the first count (first capture), field teams were instructed to go to their respective zones, previously assigned. The team in each zone walked around various locations. When they encountered a SW (verified by the volunteers) she was approached by the volunteer and asked whether someone had approached her and given her a card (she was also shown the card that the team carried). If the SW mentioned that she did not receive any card, she was given a card. This continued until about 3:00 am. At the end of this first count, the total number of cards distributed was determined for each zone (total number of street-based SW captured – first capture).

Step 13:
The same process was repeated a week later. The field teams went to their respective zones at the same time and stayed from 7:00 pm until 3:00 am. However, this time the team carried a different card. The new cards had a different colour with different information and a different sequence of numbers printed on them. Each field team was given a set number of these new cards and instructions on how to maintain records of new capture and recapture. When they encountered a SW (verified by the volunteers) she was approached by one of the volunteers and asked whether she had received a card. If she responded that she had, she was shown the new card to determine whether she received the new card or had received a card during the first count but not during this second count. If she mentioned that she received a card (different from the new one) during the previous week, she was given the new card and recorded as a recapture. If she mentioned that she never received any card, then she was given the new card and recorded as a capture during the second count. This process continued until 3:00 am.

Step 14:
Based on the total number of cards distributed during the first count (first capture – NC1), total number of cards distributed during the second count (second capture – NC2), and the total number of SW who received cards during the first count and also during the second count (recapture), calculation was conducted to estimate the total size of the street-based SW who operated primarily in the evening between 7:00 pm and 3:00 am in the selected geographical area.

Strengths and limitations

A simple two-sample CRC is relatively easy to use. It also does not require statistical expertise. However, the four assumptions listed above can sometimes be difficult to meet, and such limitations should be documented and later presented together with the results of the size estimation. Steps can be taken to avoid violating the assumptions and minimize bias that may result when assumptions are not fully met *(2)*, including by adding a third source and effectively conducting a capture-recapture-recapture. Three-source capture-recapture lets investigators relax the assumption of independent sources by using interaction terms between dependent sources in a log-linear model *(11)*.

A-11.1.3 Network scale-up method

The network scale-up method (NSUM) is used as part of a general population survey. NSUM facilitates the estimation of a participant's personal network size (i.e. the number of people personally known to the participant – the *denominator*) and probes the number of key or priority population members personally known to

the participant (the *numerator*). Dividing the numerator by the denominator yields the key or priority population prevalence within the respondent's personal network. Averaging these prevalence estimates across all survey respondents and weighting them will yield an overall estimate of prevalence in the key or priority population *(12-14)*. NSUM can also provide regional or city-level size estimates if the sample size is large enough at these levels.

Strengths and limitations

Few extra resources are required for each additional population added to an NSUM exercise, the primary resource being staff time for the additional 15–30 questions asked of each participant. When added to an already planned general population survey, the costs to produce NSUM-based estimates are small, although this process does require preparatory work. NSUM may underestimate the population size if few people know that individuals in their social network are key population members (i.e. transmission error) or when participants do not reveal that they know key population members. Transmission error potentially has the biggest impact on results. This error can be adjusted using the game of contacts (GoC) technique *(15)*, which comprises a set number of questions that are used in interviews with key population members as part of a key population BBS. It facilitates a quantitative estimate of that population's social visibility, ranging as a proportion between 0 (total invisibility) and 1 (total visibility). For example, if men who have sex with men (MSM) on average have a visibility of 0.5 (i.e. half of their network members know that a given man is an MSM), then the NSUM population size estimate needs to be doubled to account for this bias.

A-11.2 Emerging methods

In addition to the two methods discussed above, there are a number of emerging methods to use in conjunction with BBS among key population surveys.

A-11.2.1 Reverse tracking method

The reverse tracking method *(16)* is a new population size estimation method that can be integrated within a BBS. It is an indirect size estimation method that modifies or corrects the estimated total size in the sampling frame based on actual observations. The method is based on a simple mathematical principle and does not require collection of additional information.

How reverse tracking works

- the BBS sampling frame for any at-risk population group may include an exhaustive list of the sites where members of the population can be accessed for an interview and an "estimated size" for each site, based on key informant interviews during mapping or sampling frame development exercises.
- a proportion of all the sites from the sampling frame for the actual survey is selected, with the sites chosen in such a way that they are representative of all the sites in the sampling frame. For example, 30% of all the sites in the sampling frame is selected.
- during the BBS, the survey team visits the selected sites and interviews (and collects biological specimens from) a random sample of the population group at each selected site. At each selected site, information is counted and recorded; for example, the total number of completed interviews, not completed interviews, refusals and total population attached to the site ("observed size"). This information is collected through cluster information sheets for weighting purposes.
- this provides, for 30% of the sites, the "estimated size" from the sampling frame and "observed size" from the cluster information sheets of the BBS.
- the method compares the "observed size" with the "estimated size" for each selected site, calculates an adjustment or correction factor, and modifies the total estimated size in the sampling frame accordingly.

The simple mathematical formula is *(17)*:

$$N = \frac{1}{n} \sum_{i=1}^{n} \frac{Yi}{Mi} * M$$

where

N = estimate of total population size

n = total number of selected sites for BBS

Yi = "observed value" at i-th site

Mi = "estimated value" at i-th site

M = total estimated size in the sampling frame; i.e. sum of Mi for all the sites in sampling frame.

In the formula, the total estimated size in the sampling frame (i.e. M) is adjusted or corrected by

$$\frac{1}{n} \sum_{i=1}^{n} \frac{Yi}{Mi} \frac{1}{n} \sum_{i=1}^{n} \frac{Yi}{Mi}$$

and it is the average variance of the observed values from the estimated values across the selected sites.

Assumptions

There are three key assumptions that need to be fulfilled for the reverse tracking method:

- the sampling frame should include an exhaustive list of sites;

- each site in the sampling frame should have an estimated size based on key informant interviews; and
- the observed size should be recorded and be available for each selected site for the BBS.

Strengths and limitations

The strength of this method is that it does not involve any additional data collection. It uses already collected data during a BBS to generate another estimated size. The estimates from this method would be as good as the sampling frame of the BBS (in terms of its coverage of the sites). This method would not be able to estimate the hidden part of the population (i.e. that portion that never accesses these sites).

Accuracy of the "observed size" at site level during the survey is also a critical factor. In many countries where the laws are extremely harmful to target populations, it will be difficult to find sites. Another limitation is that this method could give an inflated size for a "floating population" due to presence of the same set of individuals across different time-location sampling (TLS) surveys for a particular site, and a further round of correction may be needed to address this particular issue.

A-11.2.2 Successive sampling population size estimation

Handcock, Gile and Mar *(18)* proposed a Bayesian approach to estimate population size of target populations using data from a single respondent-driven sampling (RDS) survey. In the Bayesian approach, a probability model is defined for the observed data given the population size, where it is assumed that we already have some knowledge about the current population size.

The probability model or likelihood for the data is a function of the observed personal network sizes (i.e. degrees) in the RDS sample, as well as investigators' knowledge about the population size and the unobserved network sizes, both of which are assumed to be generated from a super-population model based on some unknown distribution. The likelihood is formulated as an approximate RDS sample obtained by successive sampling from the full population with degrees dependent on a super-population parameter.

Through use of Bayes' rule, the likelihood is multiplied by an assumed prior distribution for the population size and assumed prior distribution for the super-population parameter to arrive at a posterior distribution for population size, from which it is possible to calculate a posterior mean (i.e. population size estimate) and 95% probability interval for the population size. Because prior knowledge about the size of the population might be incorrect, it is possible to perform sensitivity analyses for unknown population sizes through use of different prior distributions that may incorporate previous or concomitant information about the population size.

Handcock, Fellows and Gile *(17)* developed RDS analyst software to implement this approach. The strength of this analysis approach is that when RDS surveys are conducted, network size information is also collected and used in the estimation procedure. However, the method depends on prior information on the population size or sample proportion, and this information may be unavailable or poorly estimated, resulting in large interval estimates. In addition, the method is relatively new and will benefit from validation studies, comparison with other approaches and possible refinement. Lastly, if the survey is based on sampling designs other than RDS, then this method will not be suitable. This new approach is still being validated and may need further refinement.

A-11.2.3 One-sample capture-recapture method

Recently, Dombrowski et al. *(19)* innovatively used CRC entirely within a BBS. They used a network-based variant of CRC, whereby participants maintained anonymity. It required sampling from each participant's network connections, and matching those connections against the other participants in the sample and also the list of their respective contacts. RDS was used to recruit the first capture using the internet (Craigslist), and the second capture was based on respondents' friends in their personal networks. Recapture sampling was based on links to other participants derived from demographic and "telefunken" matching procedures – the latter being an anonymized version of telephone numbers. This technique made it possible to estimate the population size without physically recruiting a second sample.

A-11.3 Selecting a method

A-11.3.1 Factors to consider – single method versus multiple methods

All methods have limitations. Estimates are uncertain and different methods are likely to give different results. Factors that can make estimates substantially uncertain include variability in quality of service data (multiplier method), assumptions that are hard to meet (CRC) and transmission error (NSUM). There is no reason to use only a single method for estimating the size of the population. Multiple methods make for checks and balances and cross-validation. Different methods may produce different estimates. However, when different methods are used and produce estimates in the same overall range, they are likely to inspire more confidence than a single estimate produced by a single method. There are several methods

that could be integrated within any BBS. The choice of method and the number of methods used is often determined by the resources available.

Those planning and implementing size estimation activities should look into the possibility of using more than one method when planning BBS or when other surveys are being planned or implemented. Formative assessment should be conducted before implementing a BBS. Formative assessment may include mapping (social and geographical) and ethnographic observation. During mapping and ethnographic observations, attempts can be made to estimate the size of the population accessible at the sites (see Chapter A-4). If a household survey is planned, the investigators should look into the possibility of using NSUM for estimating the size of multiple key populations.

The most important step is better planning and preparation before integrating any particular method. When planning a BBS, it is important to inventory capacities as well as resources available to see which method can be integrated into the BBS. If using a service multiplier, it is important to assess what type of service data are available, the quality of the data, access to data, and any legal or other constraints to using the data. A formative assessment that includes members of the key population conducted before designing the survey clarifies the social and geographical distribution of the target population and the likelihood that the survey staff can reach them. Working with target population members informs the selection of the most appropriate methods. Piloting the method(s) in a subset of the population and validating the method(s) in a known setting (e.g. a university), where possible, are important.

The estimates are limited to the local area where the BBS was implemented. Local estimates cannot be considered as the national estimates. Considerations should also be given to sampling methods that are used for conducting the BBS; a multiplier method may be suitable to use with an RDS survey but CRC may not be. Similarly, TLS may require use of different size estimation methods to those that can be used with RDS surveys.

Estimating population size is important and useful; however, the calculated estimates should be accompanied by the corresponding degrees of uncertainty. Collected data have various limitations. Because population is an estimated number, a range is acceptable. A full description of how to analyse data from population size estimation activities is beyond the scope of these guidelines. Investigators planning to conduct a population size estimation should consult with experts and review the literature for more information.

A-11.3.2 Selecting the method for key population size estimation

This section provides examples of considerations about the choice of method to estimate key population sizes, focusing primarily on local level size estimation and on methods within the context of BBS.

Method choice considerations

Review of past survey and size estimations and key informant interviews

Planning a BBS may involve conducting formative assessment, which includes review of available data as well as focus groups and key informant interviews with the stakeholders. Any available data should be reviewed to understand whether and where target population members usually congregate and how mobile they are, perhaps due to police activity, and whether they access services. This brief preliminary information can be obtained from reviews of existing information, talking with other "stakeholders", and conducting a brief "ethnographic" assessment. The information collected through this initial "assessment" can help in understanding the feasibility of reaching the population directly – how accessible they are in general and at select locations. A formative assessment may also indicate that a BBS is not warranted at this time but that size estimation is needed.

Social visibility and mobility of key population

If the target population is mostly visible and can be accessed in certain locations within the city, then census or enumeration may be appropriate methods. Both may provide information for the creation of a sampling frame for surveys using cluster sampling. Target populations that are highly mobile may make census and enumeration more difficult and thus tend to violate CRC assumptions.

Geographical level for which the size estimation is needed

Size estimation in conjunction with a BBS can provide only local size estimates; for national level estimates, data triangulation, projection or other methods such as NSUM are needed. Whenever a household survey is being planned to be conducted, attempts should be made to use NSUM to estimate the key population sizes for the national and subnational level. NSUM may also be used within local general population surveys, especially when other methods such as CRC or multiplier may be difficult to implement.

Financial and technical resources

Census and enumeration are usually resource intensive, depending on the size of the town and the number of venues. In contrast, using Bayesian statistics on already collected RDS data usually only requires minimal

additional time during data analysis, but does require a technical understanding of the applied data analysis. CRC may require relatively more resources compared to census and enumeration and technical capacity of the staff.

Service provision
Where services tailored to the target population are available or where service providers can reliably count target population members among their clientele, this facilitates the use of the multiplier method.

Method-specific considerations

Capture-recapture method
CRC may require relatively more resources than census and enumeration, and greater technical capacity of the staff. CRC relies on several assumptions being met if it is to produce valid results. It also works best where the target population is accessible and visible (to those conducting CRC). It can be conducted as part of a survey (the survey can be used as the recapture stage) or independent from a survey.

Multiplier method
Use the multiplier method when reliable service data are available; that is, where service providers can provide accurate counts of target population members served and where the characteristics of the clientele served reflects the eligibility criteria in the BBS.

Census and enumeration
Census and enumeration should generally be seen as providing minimum key population size estimates (i.e. they often lead to underestimates). Because these methods are typically venue based they may work reasonably well only when all or most of the target population can be counted at accessible venues.

Table A-11.2 Factors to consider when choosing a key population size estimation method

	Capture-recapture	Multiplier	Census and enumeration
Target population social visibility and mobility	High mobility will bias CRC size estimate. Target population members need to be visible to capturing staff.	Less affected mobility if service data stem from same time as sampling period.	Target population members need to be visible to enumerators. High mobility may affect estimates.
Financial and technical resources	Outreach may be more resource intensive than enumeration. It may be technically demanding as assumptions for CRC need to be met.	Relatively easy. Not resource intensive.	Relatively easy. Outreach may be resource intensive, especially for census.
Service provision	Not applicable.	Accurate service provision data essential (when using service multiplier).	Not applicable.
Sampling design	May inform sampling frame for both cluster sampling surveys and RDS compatible with CRC.	Warrants probability-based survey sample.	Both may inform sampling frame for cluster sampling and RDS compatible with CRC.

CRC, capture-recapture method; RDS, respondent-driven sampling

A-11.4 References

1 UNAIDS/WHO Working Group on HIV/AIDS/STI Surveillance. Estimating the size of populations at risk for HIV: issues and methods. Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2003 (http://data.unaids.org/publications/external-documents/estimatingpopsizes_en.pdf accessed 4 August 2016).

2 UNAIDS/WHO Working Group on Global HIV/AIDS/STI Surveillance. Guidelines on estimating the size of populations most at risk to HIV. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO); 2010 (http://www.unaids.org/sites/default/files/media_asset/2011_Estimating_Populations_en_0.pdf, accessed 4 August 2016).

3 Abdul-Quader AS, Baughman AL, Hladik W. Estimating the size of key populations: current status and future possibilities. Curr Opin HIV AIDS. 2014;9(2):107–114.

4 Black JF, McLarty D, Mtasiwa DE. Capture-recapture techniques. Difficult to use in developing countries. BMJ. 1994;308(6927):531.

5 Larson A, Stevens A, Wardlaw G. Indirect estimates of 'hidden' populations: capture-recapture methods to estimate the numbers of heroin users in the Australian Capital Territory. Soc Sci Med. 1994;39(6):823–831.

6 Mastro TD, Kitayaporn D, Weniger BG, Vanichseni S, Laosunthorn V, Uneklabh T et al. Estimating the number of HIV-infected injection drug users in Bangkok: a capture–recapture method. Am J Public Health. 1994;84(7):1094–1099.

7 Neugebauer R, Wittes J. Voluntary and involuntary capture-recapture samples – problems in the estimation of hidden and elusive populations. Am J Public Health. 1994;84(7):1068–1069.

8 Sarkar S, Durandin F, Quader AA. Estimation of number of street-based female commerical sex workers (CSWS) and their HIV situation in the city of Dhaka, Proceedings of the 4th International Congress on AIDS in Asia and the Pacific, Manila, Philippines. 1997:184.

9 Stephen C. Capture–recapture methods in epidemiological studies. Infect Control Hosp Epidemiol. 1996;17(4):262–266.

10 Weir S, Wilson D, Smith PJ, Schoenbach VJ, Thomas JC, Lamptey PR et al. Assessment of a capture-recapture method for estimating the size of the female sex worker population in Bulawayo, Zimbabwe. Chapel Hill: Carolina Population Center, University of North Carolina; 2003 (http://pdf.usaid.gov/pdf_docs/PNADA379.pdf, accessed 4 August 2016).

11 Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev. 1995;17(2):243–264.

12 Salganik MJ, Fazito D, Bertoni N, Abdo AH, Mello MB, Bastos FI. Assessing network scale-up estimates for groups most at risk of HIV/AIDS: evidence from a multiple-method study of heavy drug users in Curitiba, Brazil. Am J Epidemiol. 2011;174(10):1190–1196.

13 Feehan DM, Umubyeyi A, Mahy M, Hladik W, Salganik MJ. Quantity versus quality: a survey experiment to improve the network scale-up method. Am J Epidemiol. 2016;183(8):747–757.

14 Feehan DM, Salganik MJ. Generalizing the network scale-up method: a new estimator for the size of hidden populations. Stat.AP (arXiv:1404.4009v3). 2015;(https://arxiv.org/pdf/1404.4009.pdf, accessed 8 August 2016).

15 Salganik MJ, Mello MB, Abdo AH, Bertoni N, Fazito D, Bastos FI. The game of contacts: estimating the social visibility of groups. Soc Networks. 2011;33(1):70–78.

16 Vadivoo S, Gupte MD, Adhikary R, Kohli A, Kangusamy B, Joshua V et al. Appropriateness and execution challenges of three formal size estimation methods for high-risk populations in India. AIDS. 2008;22:S137–S148.

17 Handcock MS, Fellows IE, Gile KJ. RDS Analyst: software for the analysis of respondent-driven sampling data, Version 0.42, 2013. Available from: http://hpmrg.org.

18 Handcock MS, Gile KJ, Mar CM. Estimating the size of populations at high risk for HIV using respondent-driven sampling data. Biometrics. 2015;71(1):258–266.

19 Dombrowski K, Khan B, Wendel T, McLean K, Misshula E, Curtis R. Estimating the size of the methamphetamine-using population in New York City using network sampling techniques. Adv Appl Sociol. 2012;2(4):245–252.

12. Supplemental studies

Investigators and programme implementers often need more information than that which a standard biobehavioural survey (BBS) provides. Supplemental studies can provide additional critical information to explore the individual, socioeconomic and cultural context that determines who continues to be vulnerable to HIV infection and why.

Key terms

Cohort study: An epidemiological study that follows a group of people over time to observe which and how many experience certain outcomes (e.g. HIV seroconversion). Also known as longitudinal study.

Data saturation: In qualitative research, the point at which no new useful information emerges.

Exposure: The condition of being (potentially) in contact with something, such as an infectious agent or an intervention.

Homogeneity: In respondent-driven sampling surveys, the degree to which people in a population have similar or identical characteristics.

Outcome: The measure of interest or endpoint in clinical trials or in data analysis, or the result or consequence of a public health intervention (e.g. condom use or HIV status).

Partner survey: Use of one population to recruit another related population; for example, when SW in a survey are asked to refer their clients for survey participation.

Primary survey: The survey that directly samples the target population.

Probe: Request(s) by the interviewer for more information, or a method used during an interview to help a participant recall information.

Semistructured interviews: Using an interview guide with themes or open-ended questions (grouped by topic areas) to conduct an in-depth, open-ended conversation with the respondent(s).

The BBS itself may also raise additional questions that need further investigation, including about the sex partners of the target population. The next three sections describe three data-collection approaches that can be paired with a BBS: qualitative studies, cohort studies and partner surveys.

A-12.1 Qualitative studies

Although a BBS provides information on the behaviours of participants, a qualitative study helps investigators understand the behaviours people engage in and the meanings attached to them *(1)*. This section discusses the role of qualitative studies in a BBS and introduces qualitative research methods. Specifically, it discusses two kinds of qualitative studies: formative assessment and supplemental qualitative studies. A formative assessment is conducted before a BBS and provides investigators with the information needed to plan a BBS. A supplemental qualitative study can be conducted during or after a BBS. More information on formative assessment can be found in Chapter A-4.

While quantitative research focuses on the *what, where* and *when*, qualitative research focuses on the *why* and *how*. Together, quantitative and qualitative methods can provide a deeper and more nuanced understanding of knowledge, behaviours and practices, including how individuals and communities understand HIV, sex and drug use, and how best to design interventions. Table A-12.1 provides examples of different types of information that can be obtained from men who have sex with men (MSM) and people who inject drugs (PWID) from both quantitative and qualitative interviews.

Table A-12.1 Example of information provided in quantitative and qualitative interviews

Quantitative		Qualitative	
Question	Answer	Question	Answer
For FSW: For how many years have you been selling sex for money?	3	For FSW: For how long have you been selling sex?	"I first got paid to have sex 3 years ago. A guy at a bar offered me money. I didn't feel good about it and said I'd never do it again. About a year later, my sister was in an accident and we needed money for her medical bills. Selling sex was the fastest way to make money. I stopped once we had enough money but after a few months decided to quit my normal job and do this instead because I could make so much money."
For MSM: Do you always use a condom when having anal sex?	No	For MSM: In what circumstances do you not wear a condom during anal sex?	"My boyfriend and I do not use condoms. We have been together for 2 years and get tested for HIV together. We agreed to be monogamous. I do not think he has sex with other men but it's possible. I am always the insertive partner though so my risk of HIV is less. Because of this I have less to worry about if he cheats so I am okay not using a condom with him."
For TG: Do you consider yourself as male, female, transgender or other?	Female	For TG: Tell me about your gender identity.	"I'm a woman. Some people say I'm trans and others say I'm a transwoman. But look at me, I'm a woman. I've always been a woman. It's just that I started out with male genitals. So how can I be a transwoman if I've always been a woman? It's what I am now. A woman."
For PWID: In the past 6 months when you injected, how often did you use a new, sterile needle?	25% of the time	For PWID: Why do you not always use a sterile needle?	"I used to use clean needles but one of the places I got them from closed. I don't have a steady job and cannot afford transport to the other place. It takes too long to get there too. Sometimes I reuse my own needles. Sometimes I share with others. It depends on the situation. If I'm out at a party and I'm able to get heroin, I'll use whatever equipment is there. I don't carry that stuff around with me all the time. But if I'm at home by myself, that's where I have my stash. I don't want to share it with others. I want to keep it clean."

Strengths and limitations of qualitative research methods

Like quantitative methods, qualitative methods have strengths and limitations for both the participant and study investigator, as shown in Table A-12.2.

Table A-12.2 Strengths and limitations of qualitative studies

Strengths for participants	Strengths for researchers
• Offers an opportunity to tell their story • Provides option to disclose as much or as little information as they like • Allows them to emphasize what is important to them • Can be liberating to discuss openly with someone issues that are not openly discussed in the community • Usually allows them control over the date, time and place of the interview	• Easily administered and requires very few resources to conduct • Can provide insights into quantitative data • Offers freedom to change the order of questions • Provides opportunity to prompt for more in-depth information • Provides opportunity to record spontaneous answers or responses (e.g. crying) • Provides opportunities to clarify questions • Allows spontaneous questions to be asked to address emerging issues • Provides control over who participates in interviews
Limitations for participants	**Limitations for researchers**
• Less anonymity • A participant may not feel comfortable with the interviewer or the questions • Time consuming • Emotionally exhausting	• Time consuming to analyse • Possible influence of interviewer on participants' responses • Relies on the skill of the interviewer to elicit a good interview (e.g. building rapport, providing good follow-up questions on the spot, avoiding leading questions) • Emotionally and physically exhausting to conduct interviews, particularly those that address sensitive issues such as HIV, sexuality and violence • Usually small, nonrepresentative sample; cannot apply statistical methods in analysis

Qualitative data-collection methods

As described in Chapter A-4, three qualitative data-collection methods are commonly used with BBS:

- semistructured interviews;
- focus groups; and
- ethnographic mapping.

These methods are discussed below.

Semistructured interviews

Semistructured interviews are conducted with a series of open-ended questions. Having questions that are open-ended rather than more structured (e.g. multiple choice questions) encourages two-way communication between the interviewer and the participant. It allows the interviewer to modify the course of the interview based on the participant's responses and ask follow-up questions about ideas needing further elaboration. Also, semistructured interviews provide an opportunity for investigators to delve deeply into a topic and understand the reasoning and motivation behind behaviours.

Key informants

Semistructured interviews are often used to illicit information from key informants – people who have important information that may help in designing the BBS. They include members of the target population and those with whom the members associate. Potential nontarget population key informants are listed in Table A-12.3. They can be allies in working with the target population and maneuvering local politics.

Table A-12.3 Target population and relevant key informants

Target population	Nontarget population key informants
Sex workers	Venue owners or staff, pimps, taxi drivers, HIV service providers, ministry of health officials, clients, NGO staff
Men who have sex with men	Venue owners or staff, HIV service providers, ministry of health officials, NGO staff
Transgender persons	Venue owners or staff, HIV service providers, ministry of health officials, NGO staff
People who inject drugs	Drug treatment providers, harm-reduction service providers, HIV service providers, drug dealers, law enforcement, ministry of health officials, NGO staff

NGO, nongovernmental organization

Key informants should be diverse to capture a broad range of viewpoints. For example, for a BBS among MSM, consider including MSM who are different in terms of age, neighbourhood of residence, socioeconomic status, education, HIV status, marital status, sexual identity, or the extent to which they hide the fact that they have sex with men. Many countries have multiple terms for different types of MSM. Be sure that these typologies are also represented among qualitative study participants. Similarly, non-MSM key informants should include individuals from diverse venues or service providers. If the formative assessment, BBS and supplemental studies are to be successful it is important to build trust among the population through key contacts and by visiting venues the population frequents. This also helps to mobilize the target population and introduce the survey team to them. An interview guide for use with gatekeepers can be found in Appendix I-6.

Focus groups

Focus groups are semistructured group conversations that are usually conducted with about 6–8 members of the target population under the guidance of a facilitator using an interview guide. They are useful for collecting many viewpoints quickly, particularly about topics that do not require sharing private information. Key informant interviews tend to elicit more thoughtful and personal responses, whereas focus groups tend to elicit the "public" response, and to generate group discussion and debate. For instance, participants might be more likely to discuss their risk behaviours in an individual interview than in a focus group, but might have no preference as to whether to discuss access to health services in an interview or a focus group. Using semistructured key informant interviews and focus groups together speeds data collection and enables investigators to compare responses from both settings. Focus groups also allow participants to interact with each other to spur ideas and confirm or contradict information shared by others, and collect data faster than individual interviews. It is often good to stagger interviews and focus groups so that different kinds of data from different sources can be analysed throughout the formative assessment or supplemental qualitative study.

Whereas interviews are often conducted with members and non-members of the target population alike, focus groups are usually conducted with only the target population. Separate focus groups may be needed for subgroups of a population to understand their unique perspectives. Possible subgroups might include brothel-based and street-based sex workers (SW), younger and older MSM, or methamphetamine and heroin injectors. A brief demographic questionnaire to characterize formative assessment participants can be found in Appendix I-5. A sample interview guide for semistructured interviews and focus groups can be found in Appendix I-7.

Considerations when conducting a semistructured interview or focus group

Type of interviewer

In some settings it is useful to include members of the target population in the survey team because they can facilitate buy-in and linkages to the population. These people may be seen as more trustworthy and may provide insight during data analysis. Conversely, it is possible that participants may not be as honest with their peers, especially with a person already known to them. There may also be political sensitivities to including individuals from one part of the population on the survey team and not another.

Interviewers should be trained to collect and manage data, and should use a consistent approach with a clear set of instructions. Limiting the number of interviewers results in more consistent data collection. Interviewers for the qualitative study should be given a clear introduction to the objectives of the BBS and the qualitative study, as well as a thorough review of the data-collection items and instruments.

It is important to consider the cultural and gender aspects of interviewer selection. For example, in an MSM survey, it may be useful to ask local experts (e.g. members of an MSM service organization) what kinds of people MSM will be comfortable speaking to. It may be appropriate in some settings to have peer-matched or gender-matched interviewers to put participants at ease. In other settings, participants may not accept peers as interviewers.

Audio recording interviews and focus groups

It can be difficult and distracting for an interviewer to take notes and conduct an interview at the same time because semistructured interviews are open ended, allowing for detailed and in-depth discussions of issues. It is thus recommended to audio record the interview and to have two people on each data-collection team: an interviewer and a note taker. Notes can be used to formulate follow-up questions and probes. Many ethical review committees require that consent be obtained for audio recording interviews and focus groups, and that consent be audio recorded. Audio recordings are considered personal identifiable information and require special security safeguards. The protocol should explain how digital recordings will be kept secure and destroyed. Example consent forms for qualitative interviews during the formative phase and the BBS phase can be found in Appendices I-25 and I-26.

Interview guide

An interview guide for a qualitative study may focus on a few specific topics or cover a range of topics. A qualitative interview produces the richest results when it is conducted in a conversational rather than a strict question–answer format. This allows for flexibility in the ordering of questions, the prompting of participants and, if necessary, changes in thematic emphasis based on the interests and direction of the participant and investigator. The narrative story elicited through this process results in a vivid description of the context in which participants exist and highlights the greatest influences on their environment. Administering a qualitative interview is more challenging than a quantitative interview because it requires flexibility and readiness to adapt the discussion based on the respondents' answers. Staff skilled in qualitative methods are required to conduct the qualitative interviews. Additional considerations for conducting qualitative interviews can be found in Chapter A-4.

Appendix I-24 includes an interview guide to help survey implementers design a data-collection tool for qualitative studies. It is important to remember that the interview guide is simply a guide. Survey implementers should consider the following when developing an interview guide:

- stay consistent with the goals of the qualitative survey.
- focus on a few key topics.
- keep the interview to no more than 1 hour to avoid participant fatigue.
- establish trust with the participant before the interview begins. Participants should know that they have something important to tell the interviewer, and that the interview is confidential and safe.
- use open-ended questions and limit the use of questions that can be answered with one word.
- keep questions neutral to avoid asking leading or judgemental questions. For example, ask "What do you think about HIV testing?" rather than "Don't you think you should get an HIV test?"
- include probes – requests by the interviewer for more information – on key questions, as respondents may not always provide enough information when answering a question. More information on probing is included later in this section.
- use the same terminology as participants. For example, if a FSW describes herself and other sex workers as "ladies" rather than "sex workers", the interviewer should similarly refer to them as "ladies".

Types of qualitative questions

Many types of questions may be used in a qualitative study *(2)*. The most common and important questions that can be used as part of a qualitative study include values and opinions, knowledge and information, behaviours and experiences, and feelings and meanings. Table A-12.4 provides examples of how these question types may be used to obtain information on a topic (using condoms as the topic). Additional examples can be found in Appendices I-6 and I-7.

Table A-12.4 Question types and examples about condoms

Type of question	Examples of questions
Values and opinions: These questions are aimed at understanding personal, cultural and religious values and their impact on how people think about an issue, experience or event; for example, MSM can be asked how stigma against homosexuality affects their lives and their social and sexual networks.	• Describe how effective you think condoms are. • Tell me about what young people think about condoms, especially those who use them and those who do not. • Describe how prisoners access condoms. • What do you think about prisoners having access to condoms?
Knowledge and information: These questions aim to understand what people think is factual; for example, if antiretroviral therapy is effective for treating HIV.	• Describe the places where you can access free condoms in your community. • I'm interested in learning your perspective on the relationship between condoms and HIV prevention. Do you think that condoms can prevent HIV? Why or why not? • Describe the range in cost of condoms in this town. What do condoms cost here? • Tell me about what religious groups say about condoms.
Behaviours and experiences: These questions are used to elicit people's descriptions of events, people, experiences and places; for example, the experience of an MSM telling a health service provider that he has sex with men.	• Can you tell me about how you first learned to use condoms? Who taught you? Describe for me how to use a condom. What questions do you still have? • How do you negotiate the use of condoms with casual partners? Tell me about the last time you negotiated condom use with a casual partner. How did it go? Is that how you usually approach it? • Who do you use condoms with and why? Tell me about your partners who you use condoms with and your partners who you don't use condoms with. How do you make the decision about when to use a condom? • When are you most likely to use a condom? • When are you least likely to use a condom?
Feelings and meanings: These questions are used to understand the emotional response of a participant (or others) in response to an issue or an event; for example, how a person felt after being diagnosed with HIV.	• Describe your feelings when you ask your sex partner to use a condom. How does your sex partner usually react? What feelings come to you when your sex partner reacts this way? • How do you feel when your sex partner refuses to use a condom? • Could you tell me what it's like when your sex partner asks to use a condom? How do you respond? • What is it like going to buy condoms at a kiosk or pharmacy or requesting them from a health provider?

Probing

Beyond knowing the themes of an interview and the type of questions to ask, a good interviewer should know how to probe for more information. Probes can ask for elaboration, a definition, a comparison or context *(3)*. The ability to probe is one of the most important skills an interviewer can have. The interviewer must ensure that probes are respectful and nonjudgemental, and draw the participants into a conversation. The most interesting stories or information may sometimes be about events or opinions that the participant is uncomfortable talking about. In these cases, the interviewer needs to have already established trust with the participant, and may need to offer a reminder of the importance of the participant's story at that moment. The interviewer must also know when to probe. To do this, the interviewer must understand the goals of the qualitative study and be able to recognize immediately when the participant has mentioned something that could be an important finding. Six useful probes are as follows *(4)*:

1. Elaboration (to obtain more detail):

   a. Can you tell me a little more about that?
   b. What else did she say to you?

2 Continuation (to encourage the participant to keep talking):

a. Please go on.
b. What happened then?

3. Clarification:

a. I am not sure I understand what you mean by that.
b. What do you mean by that?
c. Do you mean that you saw her do that?

4. Attention:

a. That is really interesting. Can you please tell me more?
b. I see.

5. Completion:

a. You said you spoke to him. What happened then?

6. Evidence (to identify how sure a person is of their interpretation):

a. How certain are you that things happened in that order?
b. How likely is it that you might change your opinion on that?

Avoid asking "why" when probing because it can feel confrontational. Instead, use probes such as "describe", "how come" or "tell me more about …". Participants can provide information in a qualitative study through more than just words; for example, they can draw a timeline of their sex partners in the past 6 months and indicate how they met them and which partners know each other.

Ethnographic mapping

Ethnographic mapping uses simple graphics or maps to convey information about the environment or context in which a survey is being conducted. It helps investigators to understand where the target population congregates and the activities they engage in. Maps can show where risk behaviours occur (e.g. where sex is sold or drugs are used), and locations where the target population gathers (including expected numbers by location, time and subgroup) or accesses health services; thus, maps can show where the target population can be approached.

Ethnographic mapping can indicate whether time-location sampling (TLS) is the most appropriate or feasible sampling method, based on whether populations can be readily mapped and accessed. While ethnographic mapping may help with creating a list of venues for cluster-based methods, it may not include all venues and it may include items that are not relevant to a sampling frame. Mapping of resources, including HIV services, where condoms or lubricants can be obtained or where needles can be exchanged, can facilitate identification of referral facilities for the BBS and the location of the survey site.

Maps should also identify barriers that prevent people from accessing a location *(5, 6)*. For example, a religious institution or law enforcement agency near a needle and syringe exchange centre or a survey site may discourage PWID from visiting the centre or participating in the survey. Maps have the added benefit of providing important information that can be used immediately to improve or increase HIV services even before the BBS is completed.

Maps must be kept secure and shared with only certain stakeholders, including service providers and community members, to protect the target population. The map should not be too detailed about specific locations. Table A-12.5 provides examples of physical structures, geographical areas and behaviours that may be included in a map.

Table A-12.5 Examples of content of ethnographic maps

Physical structures	Geographical areas	Social behaviours
• Service providers or outreach areas • Brothels • Bars, clubs • Pharmacies • Parks • Hotels • Schools • Transport hubs • Community centres • Churches, temples or mosques • Cafes • Health clinics or hospitals • Police stations • Military barracks • Homeless shelters • NGOs • Safe houses for those escaping domestic violence • Needle and syringe exchange centres • HIV testing and counselling sites	• Ethnicity of population • Bus routes • Street intersections • Low-income areas	Places where: • sex is sold or traded • sexual partners are found • drugs are purchased • drugs are used • target population members socialize together • violence occurs • outreach services occur

NGO, nongovernmental organization

Ethnographic mapping can take different forms. Key informants or focus group participants can either draw maps themselves or add important information to existing maps. They can also be asked to verify the existence of pre-identified items on a map and make any changes. Maps from different groups or individuals can then be compared and verified. As places on a map may differ by the number or type of people who go to the location, as well as the day or time they may be there, it is useful to also solicit this information. An interview guide for collecting mapping information as part of a key informant interview or focus group can be found in Appendix I-8. For more information on mapping, see *Tools together now! 100 participatory tools to mobilize communities for HIV/AIDS (7)*.

Maps vary depending on the participants who create them. For example, an older MSM who is married with children may create a very different map of where he goes for sex or health services than a young MSM who is open about his sexuality. Their barriers to health services are also likely to differ. Neither map is right or wrong; each provides details to investigators on MSM activities in that setting. When diverse participants contribute to ethnographic mapping efforts, investigators get the most complete map.

Many mapping exercises include visits to places on the map to observe the site and confirm information provided by participants about locations and behaviours *(8, 9)*. Efforts should be made to visit "high density" and "low density" locations, because those who frequent them may have different characteristics, behaviours and attitudes.

Investigators should note whether there is homogeneity (i.e. people there are similar to one another) within a venue, and across venues. This can facilitate sampling if conventional cluster sampling (CCS) or TLS are used. Observing and engaging the community can bring insight to both the development and conduct of the survey, as well as the interpretation of the findings. While this is relevant for all sampling methods in BBS, it is most relevant to cluster-based methods (described more in Chapter A-9).

Qualitative data analysis

After the interview or focus group, the interviewer and the note taker should note their impressions about the session, its main themes, and comments and reactions of the participants. This should be followed by a debriefing with the entire survey team. The debriefing should compare the newest data with data from previous interviews and focus groups, to develop an understanding of the context and opinions of the target population.

Investigators should frequently reflect on the objectives of the formative assessment and the BBS during data analysis to maintain the focus of the analysis. Because formative assessments usually focus on a few specific questions with only a few participants, data analysis software is not necessary. However, if the number of

participants and amount of data in a supplemental qualitative study are large, the use of electronic data analysis software is recommended. Qualitative software packages allow the user to transcribe and code interviews, making it easier to identify themes and ideas. It is good practice for each interview to be coded by two people to ensure consistency and quality. One common software application is Dedoose[1]; others are NVivo[2] and ATLAS.ti[3].

If multiple people conducted the interviews, they should debrief regularly about themes and questions emerging from their interviews, to confirm common themes and explore issues in more depth in future interviews. For example, an interviewer may learn that some transgender persons (TG) indicate that they do not sell or exchange sex; however, upon further exploration the participant may admit to seeking out sex partners who will give things. The identification of this subcategory of transactional sex should be shared among the study team so that it can be explored in future interviews by the entire team.

In another example, previous interviews may indicate that PWID had few problems with police or health-care workers; then one interview participant reveals there are actually many problems. New or different information such as this should be explored in additional interviews to find out why this person's experience is so different. It could be related to where the person spends time or which other activities the person is engaged in (e.g. theft or selling sex).

Consent to participate in a qualitative study is separate from consent for the main survey and should be obtained before the start of the qualitative interview. Many institutional review boards (IRBs) consider a person's voice to be identifiable information. Investigators should check in advance what relevant ethics committees will allow investigators to collect. Audio recordings of interviews should be kept in a secure place and destroyed as soon as the interview has been transcribed and verified. Transcriptions should be verbatim (i.e. an exact reflection of the entire conversation). The transcriber should not abbreviate or paraphrase any text. Noteworthy appearances and undertones (e.g. laughing or crying) should be captured and clearly noted as remarks or comments by the transcriber.

During the course of a qualitative interview, participants may provide identifying information about themselves or other members of the target population. This information should be altered to ensure anonymity. Real names should never be transcribed or used in data analysis.

Timing of supplemental studies

Supplemental qualitative studies can be implemented during or after the BBS. In a BBS that does not collect identifiable information from participants, it is best to implement the qualitative study during the survey, so that individuals can easily be invited to participate. The qualitative interview should be conducted on another day, so that the participant does not become overwhelmed by the length of the interview process. Individuals are more likely to be available and interested in participating in the qualitative study if they are invited before the BBS ends. Participants should be provided with additional compensation for their transport to the survey site and time spent during the qualitative study.

Sampling

Participants in supplemental qualitative studies are often sampled purposively; that is, they are selected based on their responses during the quantitative interview. Common criteria for selecting participants include demographics, HIV status, treatment status, history of exchanging sex (for non-SW), and history of violence or discrimination. Other examples include MSM who have female sex partners and PWID sampled in an SW survey. In addition, it is important to select participants who are talkative and seem as though they have information to share. Alternatively, survey implementers may sample every n^{th} participant in the BBS. It is important to note that doing this does not result in representative qualitative data.

Sample size

As described in Chapter A-4 on formative assessment, sample size is not set *a priori* for qualitative research. Sample size in qualitative research is based on data saturation (also known as redundancy); that is, the point at which no new useful information emerges from the interviews. It is possible to interview 100 participants who have little to share and obtain very little information. It is also possible to interview 20 participants and obtain all the information needed. The depth of information shared by participants and the analytical abilities of investigators are more important than the sample size *(2)*. The protocol should explain the rationale for the selected sample size *(10)*.

A supplemental qualitative study will probably require a larger sample size than a formative assessment due to the greater diversity and depth in responses. Depending on the questions of interest, it may be necessary to sample more participants of a certain type (e.g. individuals who always use condoms). The exact number of semistructured interviews to conduct depends on whether focus group discussions are conducted with members of the target

[1] www.dedoose.com

[2] http://www.qsrinternational.com/product

[3] http://atlasti.com

population. Often, the more focus groups conducted, the fewer semistructured interviews needed. It is common to conduct approximately 6–9 semistructured interviews with diverse key informants who are not part of the target population, and 12–15 members of the target population. During data analysis of the qualitative interviews, investigators may decide to stop coding (analysing) interviews once the stage of saturation or redundancy has been reached.

A-12.2 Cohort studies

BBS are generally cross-sectional, meaning that information on each participant is collected at only one point in time. Cross-sectional surveys cannot show causation, or measure the effect of exposures on outcomes over time, or time to an event (e.g. death or initiation of treatment). Cohort studies, in contrast, can show causality because they follow the same participants over time *(11)*.[1] Cohort studies enable survey implementers to explore many things, including:

- directly observed HIV incidence;
- behaviour changes;
- linkages to HIV care and treatment; and
- retention in HIV care and treatment.

Despite the important information that cohort studies can provide, they are not very common due to the financial costs and logistical challenges of following the same people for an extended time. They also require substantial planning to implement. However, cohort studies can be conducted when resources allow and when it is important to demonstrate causal relationships. In these cases, a BBS can provide the initial sample for a cohort study if it is planned well in advance. Participants in the BBS can then be offered enrolment in the subsequent cohort study. This chapter provides some guidelines for conducting a cohort study as part of BBS planning.

Timing

The decision to implement a cohort study together with a BBS should be made early in the planning stages. This will allow survey investigators to assess the feasibility of a cohort study during the formative assessment (see Chapter A-4). Implementers can use that assessment to determine the willingness of target population members to:

- participate in a cohort study;
- return to the survey site for follow-up visits for additional interviews or to provide biological specimens; and
- be contacted for reminders about scheduled survey appointments.

Sample size and sampling method

A BBS generating a representative sample is a good way to start building a cohort. The calculated sample size for the BBS should be used as a starting point to determine the sample size for the cohort study. The following additional factors should be considered to determine whether the sample size should be increased due to the addition of the cohort study:

- the frequency of the event of interest:
  - This is the most important factor. If survey implementers are interested in observing HIV incidence, a relatively rare event, they will either need to follow participants for a long time and risk loss to follow-up, or follow more people for a shorter period. They should consider the cost and logistical trade-offs.
- cohort-specific enrolment criteria:
  - At time of enrolment into the cohort, participants should lack the main outcome of interest. For example, if the objective is to examine HIV incidence, only HIV-negative participants may be offered enrolment.
- unwillingness of BBS participants to participate in the cohort study:
  - Not all participants in the BBS will be willing to participate in the cohort study. Willingness to participate should be assessed during the formative assessment.
- loss of participants:
  - Participants may initially be willing to participate in the cohort study but eventually withdraw or be lost to follow-up due to a lack of time or interest in the survey. This will reduce the power to detect statistical significance as the sample size decreases; it also opens the study to bias if the participants who drop out are different from those who remain in relation to the variables of interest.

Participation in the BBS counts as the first observation (baseline) for each participant in the cohort study, which continues after the BBS is completed.

Data-collection instruments

Depending on the goal of the cohort study, investigators may need to develop additional tools and forms that are not used during the BBS. Additional tools may include appointment reminder cards, and personal identification tracking forms to contact participants before future visits.

If adapting a questionnaire that was made for a cross-sectional survey, it is useful to consider how the variables in the questionnaire track changes over time.

Data analysis

Data from a BBS and supplemental cohort study can be analysed separately or together, using the BBS data as the cohort's first observation point *(12, 13)*. BBS data should be analysed using appropriate analysis techniques for cross-sectional surveys. These data should be analysed while the cohort study is being implemented to make results available faster. Other analysis methods, namely those using person-time to measure exposures and outcomes, are used in cohort studies *(11)*.

Operational and cost considerations

Cohort studies can be expensive because of the amount of time they take to implement and the costs associated with tracking people over time. Collaborations with service providers may reduce costs but may also risk discouraging participation by individuals who do not want to be associated with these providers.

Survey implementers should ensure that budgets can sustain staff salaries and other recurring costs during the cohort's duration. Conversely, extra coordination will be necessary if participants start returning for their second observation while recruitment (first-observation visits) is still going on. This would require survey staff to manage two different procedures and possibly high participant load at the same time.

Participants are unlikely to forget about a follow-up survey visit if it is scheduled shortly after the previous visit, but too many visits may result in participant fatigue and dropout. The number of survey visits per participant and the amount of time between visits depend on the goals of the cohort study and study operations. For example, the standard indicator time frame (see Section 3) is to measure linkage to care within 6 months. Thus, asking HIV-infected participants if they have gone for care 1 month after being diagnosed does not facilitate reporting on this indicator. Similarly, if seeking to measure directly observed HIV incidence and compare it to results from an assay measure, the time between survey visits will depend strongly on the kind of HIV test being used and its window period.

Human subjects considerations

If a cohort study is planned on top of a BBS, the consent language used for the BBS will need to reference that survey staff will also approach the potential participant about enrolment in the cohort. A separate protocol and review by an IRB will be required.

A-12.3 Partner surveys

A-12.3.1 What are partner surveys?

In a partner survey, participants in a primary survey recruit another related population to produce a second sample. Participants in the primary survey may be sampled using any method, whereas participants in the partner survey are identified using referral sampling. For example, a primary survey of SW could be coupled with a partner survey of their clients. In this case, SW who participated in the primary survey would be asked to recruit some of their clients for the partner survey. Unlike snowball sampling, the eligibility criteria for the partner survey is different from that of the primary survey, and partner survey participants do not recruit others.

Investigators may consider using a partner survey when direct sampling methods such as snowball sampling, CCS, TLS or respondent-driven sampling (RDS) are unable to successfully reach the target population. Direct sampling is difficult or impossible when the target population members do not interact with one another (e.g. clients of SW), do not frequent venues (e.g. sex partners of PWID), or are unaware of their membership in the population (e.g. female partners of MSM, regular partners of SW, or individuals whose partners are in multiple concurrent partnerships). This method can be employed to recruit female sex partners of MSM, clients of SW or sex partners of TG.

The probability of being recruited into a partner survey is usually unknown. Researchers are working to identify methods for obtaining probability-based results using this method, and partner surveys using sampling weights have been published *(14)*. Currently, there are three options for analysis:

- analysing the data as though it came from an unweighted nonprobability sample;
- applying the same weights as those used that were applied to their recruiter in the primary survey; and
- deducing a weight for the participants; for example, by asking MSM how many female sex partners they have, or asking SW how many clients they have had in the past 2 weeks.

A-12.3.2 Preparing for an indirect sampling survey

Adding a partner survey to a BBS requires additional preparations:

- planning how to indirectly sample participants;
- determining the target sample size;

- protecting the confidentiality of primary survey participants from their recruits and vice versa; and
- identifying an appropriate survey site.

Planning how to indirectly sample participants

The most common means for recruiting participants for a partner survey is to give participants of the primary survey (e.g. SW) a fixed number of coupons to distribute to another target population (e.g. clients of SW). The coupons used for each population should be distinct in colour and survey identification numbers (IDs), so that participants in the partner survey are not recruited into the primary survey or vice versa. More information on the use of coupons for recruitment can be found in Section A-9.7 on RDS. Although it is not yet clear whether RDS analysis methods can be extrapolated to partner survey data, investigators should still track who recruits whom to facilitate different analyses.

The number of coupons given to participants of the primary survey to recruit participants in the partner survey will depend on the characteristics of this second population. For example, SW could be given three coupons to recruit clients because they have multiple paying clients, but only one coupon for regular nonpaying partners. It is preferable to recruit one client each from many SW rather than multiple clients from just a few SW, because the clients of any one SW may have more in common with each other than with clients of other SW. The number of coupons per secondary population may also depend on the survey budget.

Determining the target sample size

It is difficult to anticipate the sample size for the partner survey because it is often unknown how many partner survey participants a primary survey participant may recruit on average. This depends on the number of "partners" any one primary survey participant may have, the willingness of the primary survey participant to actually recruit (hand out a coupon to) a "partner", and the willingness of such partners to enrol. For the purposes of planning and writing the protocol, investigators can estimate the sample size relative to the number of participants in the primary survey. For instance, in a primary survey of MSM and a partner survey of their female sex partners, investigators may anticipate that not all MSM will have female sex partners, so the sample size of the partner survey will be smaller than the size of the primary survey. In contrast, a partner survey of clients of SW may approach that of the primary survey of SW, because all SW have clients and most have more than one. More information on sample size can be found in Chapter A-10.

Protecting the confidentiality of primary survey participants from their recruits

The ethics and confidentiality of partner surveys are complicated by the necessity to protect the confidentiality of the primary survey participants. For example, investigators should not tell female sex partners of MSM that they are eligible for the survey because their recruiter (sex partner) is an MSM. In this case, providing the participants with information about the person who recruited them is a violation of confidentiality. Similarly, participants in the partner survey should not be informed of the eligibility criteria for either survey population (primary and partner survey). For example, in a survey examining sexual networks and multiple concurrent partnerships, partners of survey participants should not learn of the existence of concurrent sex partners or even why they are eligible to participate. Preventive measures could include scheduling appointments for different participants on different days.

Survey staff should be trained to withhold information from participants of the partner survey about the true nature of the primary survey population. Recruiters may need to be trained to provide an alternate explanation to their recruits about the purpose of the survey and the eligibility criteria that still allows them to be truthful. One useful way of explaining to partner survey participants why they did not receive coupons to recruit others is that a lottery system was used to select who was asked to do this.

Identifying an appropriate survey site

Although the primary survey may be implemented at one or more fixed sites or at venues, a partner survey is best implemented at a fixed survey site that participants can easily visit according to their schedule. It is not necessary for primary survey participants to visit this survey site. It may also be helpful to pick a site where a diverse population is present, so that no population stands out to the participants. For example, in an RDS survey of MSM who also recruit their female sex partners, the site would be frequented by many gay men but few women, which could cause suspicion among the female partners. If instead many groups use the site, MSM and their female sex partners will blend in with the other populations. Alternatively, separate days may be dedicated to enrol either primary or partner survey participants. More information on how to identify and set up survey sites can be found in Section A-9.7 on RDS.

A-12.3.3 Implementing a partner survey

The survey procedures for partner survey participants may be similar or identical to those of the primary survey. If the partner survey aims to collect biobehavioural data, then these participants may undergo similar survey activities to that of their recruits (e.g. responding to questionnaires and providing biological specimen). But survey activities may also differ. Investigators may wish, for example, to conduct qualitative rather than quantitative interviews in the partner survey, especially if investigators suspect that the achievable sample size is small.

Below are several considerations for survey procedures in the primary and partner surveys.

For participants in the primary survey:

- recruitment
  - participants in the primary survey should be asked to recruit individuals for the partner survey only after completing the survey questionnaire. This prevents participants from confusing the activities of the two surveys. It also helps investigators to assess whether the primary survey participant may recruit "partners" and how many they may recruit. For instance, only MSM who indicated in the survey that they have female sex partners should be asked to recruit them.
  - In RDS surveys, primary survey participants should first be trained in recruiting peers for the primary survey. They may then be trained to recruit for the partner survey. Investigators should use different coupons (using different colours or other visual cues) to help to keep the two processes distinct. For example, green coupons could be used for recruiting peers (primary survey) whereas orange coupons could be used for the partner survey.
  - If the primary survey sampling method is not RDS, survey staff should introduce participants to the concept of peer recruitment.
- biomarkers
  - because partner surveys must generally be conducted in a survey office, primary survey participants surveyed at a venue, as in TLS, can go to the survey office to collect any laboratory-based test results.

For participants in the partner survey:

- eligibility
  - the recruitment coupon and eligibility should be verified.
- recruitment
  - partner survey participants may not know the real reason they were recruited. In some cases it may be possible to ask partner survey participants to recruit others, but doing so may further complicate data analysis.
- biomarkers
  - because a survey site is used, participants can return to the office to collect any laboratory-based test results. Where feasible, it is best to use rapid tests to ensure that participants receive the results of their tests in real time.

The research team should track the link between recruits and recruiter through referral management software,[4] Microsoft Excel or some other system to allow for possible weighting of the survey data. More information on tracking sampling can be found in Section A-9.7.

A-12.3.4 Coordination

Surveying more than one population at a time requires substantial coordination. Separate eligibility screening, coupon systems, consent forms and data instruments may need to be used. If the protocol indicates that participants will not be provided with complete information about their recruiter or the reasons for their eligibility, survey staff should be trained to ensure that such information is not inadvertently disclosed; for example, that the female sex partner of an MSM is not accidentally given a consent form for MSM participants or coupons to recruit others. Survey staff should use different participant ID numbering systems, different coloured paper for tracking, or other defined procedures for each target population to keep them distinct.

A-12.4 References

1 Ozawa S, Pongpirul K. 10 best resources on ... mixed methods research in health systems. Health Policy Plan. 2014;29(3):323–327 (http://www.ncbi.nlm.nih.gov/pubmed/23564372, accessed 30 August 2016).

2 Patton M. Qualitative evaluation and research methods. Beverly Hills, Sage publications. 1990.

3 Mariampolski H. Qualitative market research. Sage. 2001 (http://dx.doi.org/10.4135/9781412985529.n312, accessed 30 August 2016).

4 Rubin H. Qualitative interviewing (2nd ed.): the art of hearing data. Thousand Oaks, California, 2005.

[4] www.respondentdrivensampling.org

5 Emmanuel F, Blanchard J, Zaheer HA, Reza T, Holte-McKenzie M, team H. The HIV/AIDS Surveillance Project mapping approach: an innovative approach for mapping and size estimation for groups at a higher risk of HIV in Pakistan. AIDS. 2010;24 Suppl 2:S77–84 (http://www.ncbi.nlm.nih.gov/pubmed/20610953, accessed 30 August 2016).

6 UNAIDS. Issues brief: local epidemics. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2014 (http://www.unaids.org/en/resources/documents/2014/20140707_JC2559_local-epidemics, accessed 16 November 2016).

7 K4Health. Tools together now: 100 participatory tools to mobilize communities for HIV/AIDS. 2016 (https://www.k4health.org/toolkits/pc-hivaids/tools-together-now-100-participatory-tools-mobilize-communities-hivaids, accessed 4 August 2016).

8 Aronson RE, Wallis AB, O'Campo PJ, Whitehead TL, Schafer P. Ethnographically informed community evaluation: a framework and approach for evaluating community-based initiatives. Matern Child Health J. 2007;11(2):97–109.

9 Spradley JB. Participant observation. Belmont, CA, Wadsworth Cengage Learning. 1980.

10 Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods. 2006;18(1):59–82.

11 Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet. 2002;359(9303):341–345.

12 Horyniak D, Higgs P, Jenkinson R, Degenhardt L, Stoove M, Kerr T et al. Establishing the Melbourne Injecting Drug User Cohort Study (MIX): rationale, methods, and baseline and twelve-month follow-up results. Harm Reduction Journal. 2013;10(1):11.

13 Yang H, Hao C, Huan X, Yan H, Guan W, Xu X et al. HIV incidence and associated factors in a cohort of men who have sex with men in Nanjing, China. Sex Transm Dis. 2010;37(4):208–213.

14 Shah NS, Shiraishi RW, Subhachaturas W, Anand A, Whitehead SJ, Tanpradech S et al. Bridging populations—sexual risk behaviors and hiv prevalence in clients and partners of female sex workers, Bangkok, Thailand 2007. J Urban Health. 2011;88(3):533–544.

13. Data management

The purpose of this chapter is to describe data-management planning to facilitate the collection of high-quality data. The survey team must have a clear plan for compiling and processing data, and for ensuring the confidentiality of survey participants in the earliest stages of survey planning. Investing in data management before data collection saves time and allows effective use of resources.

Key terms

Data confidentiality: The protection of data against unintended or unauthorized access.

Data security: The requirements (technical and physical) that keep data secure and confidential.

Creating a data-management plan

A data-management plan helps investigators to design and implement the survey; monitor data collection; prepare for data entry, data cleaning and data analysis; and guide data use and sharing. A data-management plan generally includes the following components:

- data documentation
- data dictionary
- unique participant identifier
- data-quality checks
- skip patterns
- data entry
- data confidentiality
- data access and use
- data backup
- data archiving and version control
- data security and storage

A-13.1 Data documentation

Data documentation is a detailed record of how the data are collected, defined, organized, processed and analysed. Both paper-based and electronic data should be documented at each step of the data-management process to ensure all data can be traced accurately from collection to analysis.

Examples of survey-level data documentation:

- description of the database: narrative of the objectives and methods of the survey to provide context for the database;
- detailed chronology of events and activities through the life of the survey (e.g. starting and ending dates of data collection, and number of participants);
- compilation of all data-collection instruments (e.g. questionnaires, consent forms, specimen tracking sheets and screening tools);
- list and descriptions of all materials used (e.g. interview schedules, logbooks, standard operating procedures [SOPs], training materials and confidentiality agreements);
- information on data validation and quality assurance procedures performed;
- information on data confidentiality, access and conditions for sharing or use;
- SOPs on how to make changes to the data (e.g. cleaning, error checking and creation of variables); and
- description of changes made to the data since its collection (e.g. cleaning, error checking and creation of variables).

A-13.2 Data dictionary

A data dictionary or codebook describes the content of a dataset. Many data analysis or data-capture systems can automatically generate a data dictionary or codebook. An example of a data dictionary is given in Appendix I-28.

A data dictionary typically contains the following information:

- variable names, labels, descriptions and values;
- acceptable values;
- variable type (e.g. text or string, numeric, data or time, or Boolean);
- codes and reasons for missing data;
- checks or restrictions (e.g. valid ranges, logic checks or skip patterns);

- derived data or variables created using code, algorithm or command file; and
- weighting variables created, if needed (for probability-based sampling methods).

A-13.3 Unique participant identifier

Each participant should be assigned a unique participant identifier or survey identification number (ID). All data collected for a participant must be linked back to this survey ID. The data-management plan should include specific quality-control measures to prevent reuse of the survey ID. Data entry or transcription errors can be reduced by using barcodes with the survey ID on preprinted labels that can be placed on forms and specimens. This requires the use of barcode scanners or other devices (e.g. tablets) with digital cameras and with the capacity to scan coded data. There should also be processes to account for unused labels to prevent accidental reuse. If barcode scanning is not feasible, preprinted labels should still be considered to reduce transcription errors and the reuse of survey IDs.

Most laboratories assign their own unique ID to each specimen. Survey investigators should determine whether they need additional data-management tools to link the survey ID and the specimen ID. Linking the main survey ID with the specimen ID can help detect coding errors (or in respondent-driven sampling [RDS], the recruit or coupon ID). Also, the specimen ID can serve as a backup unique participant identifier.

A-13.4 Data-quality checks

Assessing data quality is an integral part of data collection, data entry, data cleaning and data analysis. Quality-control measures help prevent data loss and minimize data entry errors, ultimately saving time and resources. Examples of data-quality checks include creating validation checks and skip patterns.

Ideally, validation rules for a questionnaire should be developed *before* data collection for both electronic and paper-based administration. In electronic data collection, data validation rules can be programmed into the software to automatically detect invalid responses. For paper-based data collection, interviewers or their supervisors in the field will have to detect the errors themselves. Therefore, interview staff must be properly trained to apply the validation rules. Table A-13.1 shows sample validation rules for a survey.

Table A-13.1 Validation rules

Validation rules	Application	Example
Acceptable values	Assign values that are acceptable for a categorical variable, including values that represent "Other", "Don't know" and "Refuse to answer"	The question "Have you ever had sex?" could have the following values: • Yes • No • Don't know • Refuse to answer The values "1", "2", "8" and "9" are the only acceptable values for responses to this question.
Valid range	Assign a minimum and maximum value based on allowable responses	The question "What month were you born?" has the following minimum and maximum values based on the number of months in the year. • Minimum: 1 • Maximum: 12 Values outside of this range are not allowed.
Logic or consistency checks	Ensure data collected at different points throughout the interview are internally consistent	If a participant responds he is 22 years old, all age-related responses cannot be larger than 22. For example, the participant cannot respond "27" to the question "When did you first have sex?".
Mandatory fields	Require that a response be provided. For some fields, responses such as "Don't know" and "Refuse to answer" may not be relevant as response options; for example, questions determining eligibility. In general, however, it is preferable not to leave blanks and to record a response for each field even if it is "Don't know" or "Refuse to answer".	Example 1: Survey ID Example 2: The participant's sex is needed because it will affect the type of questions asked in the questionnaire.

ID, identification number

A-13.5 Skip patterns

Skip patterns differ from validation rules. They ensure that only relevant questions are asked to each participant, because questions are only asked based on responses to previous questions within a questionnaire. For example, if a participant responded that he is male, a question regarding pregnancy will be skipped.

Incorporating skip patterns can improve the flow of administrating a questionnaire. It is important to pilot test a questionnaire to make sure that skip patterns are applied correctly.

A-13.6 Data entry

Data entry errors are common and can have a measurable impact on data analysis. Double data entry, especially for paper-based questionnaires, is strongly recommended to produce a valid dataset. Double data entry is when two different people enter the same data and compare the data for errors. Although this process costs more than single data entry, it reduces the number of data entry errors.

There are two ways to clean data after double entry:

- compare the two data entry files to find discrepant observations. This can usually be done automatically using statistical software (e.g. Epi Info, SPSS, SAS or Stata).
- have the statistical software flag inconsistencies as data are entered the second time. When an inconsistency is identified, the person responsible for entering the data the second time determines whether the original entry is correct.

Cleaning of data entry errors is further described in Chapter B-3.

A-13.7 Data confidentiality

Data confidentiality is the protection of data against unintended or unauthorized access, especially participants' identities. Biobehavioural surveys (BBS) should be done anonymously; if this is not possible, investigators must implement measures to assure data confidentiality and security. Data confidentiality must be maintained during data entry and processing, and staff should be trained in ethical issues important for surveys. All staff with access to the data should sign a confidentiality agreement before the start of the survey. This agreement outlines how data must be kept confidential and secure, and it may include information on any relevant laws that exist in the country and the penalties for breaching confidentiality. The original agreement should be held in the staff member's personnel file and a copy given to the staff member. Staff should review confidentiality and data security procedures, and sign the agreement on an established periodic basis (e.g. annually). An example of a data-use agreement is in Appendix I-37.

Any personal identifying information collected (e.g. name, ID and address) should be removed from electronic forms before data entry. Personal identifying information should only be collected if it is required (e.g. returning laboratory results). If collected electronically, this information should be encrypted. If collected by paper, the information should be stored separately from other collected data. Special consideration is needed for relational data where connections between variables in related datasets can disclose participants' identities. Sensitive words – for example, men who have sex with men (MSM), people who inject drugs (PWID) and female sex workers (FSW) – should not be used on any data-collection tools or other educational materials (e.g. educational pamphlets). The same is true for geospatial data that may reveal sensitive information. Documents such as signed consent forms should be stored separately from the data files.

A-13.8 Data access and use

Guidelines for data access and use should be an integral part of the data-management plan. This includes determining:

- data access:
  - who has access to the data;
  - who has access to computers, tablets and servers that contain data;
  - who has the authority to give access to the data;
- data ownership:
  - who owns the data;
  - whether or how the data will be shared;
- data use:
  - how the data will be used; and
  - who will manage the data.

Defining access should not be limited to datasets but also data collected on paper such as logbooks, questionnaires, checklists and laboratory results. Access to data with personal identifiers should be limited to staff who are authorized to access those data for operations (e.g. the data manager checking for duplicates or a counsellor giving a participant test results). Data analysis should be conducted only on a de-identified dataset.

A-13.9 Data backup

It is imperative that all data are protected against loss. Electronic data can be lost due to hardware or software failure, viruses, power failure or human error. Paper forms can be lost during collection or transfer, destroyed by accident, ruined by physical means (e.g. beverage spills), or misfiled and never located again. These and other topics are described in an example data-management SOP for surveys using a questionnaire development system (QDS) in Appendix I-27.

Investigators should back up data daily during data collection. Systems that automate data backups are ideal because they do not require conscious effort to safeguard data. The backup plan should include strategies for all systems where the data are stored, including tablets and network-based and non-network-based computers. Backup files should be verified and validated regularly by comparing them with the originals and checking them for completeness. The data backup plan should include all the steps for data backup and validation of the backups, and should be an appendix to the data-management plan. Access to backed-up data should be restricted, similar to access to the original data file locations.

A-13.10 Data archiving and version control

Every time a dataset is backed up a new version is created. All versions of the database must be archived and good version control practices should be implemented, to protect against loss of information. Versions of the files and backup files can be identified by dates recorded in the file name or version numbering in the file name (e.g. v1, v2, v3, and so on). For paper-based data-collection tools, the version number can be inserted as a header or footer on the document.

Maintenance of the master files of each dataset should be assigned to specific team members who are responsible for updating specific components of the datasets or database, as needed. Any changes should be managed through direct supervision according to the data-management plan. The process of freezing the database (i.e. when the data files are archived) should also be included when describing version control.

Freezing the database involves making copies at specific intervals for archiving and data use. This ensures that any analysis from the database can be reproduced and provides a structured time frame for incorporating changes.

A-13.11 Data security and storage

Storing backups solely on-site does not provide adequate protection, because the survey office's computers are usually part of the same network, or the office may be burgled, or raided by law enforcement. Furthermore, safeguards must be put in place to limit risks to participants through accidental or malicious disclosure of data by theft of forms or data-collection devices. These safeguards can include use of passwords, encryption and limiting the data stored at survey sites, especially on portable devices.

To secure paper-based data:

- keep all hardcopies (paper questionnaires and other survey forms if applicable) in a locked cabinet with controlled access;
- plan secure transport systems of paper forms to prevent loss of data;
- store personal identifying information separately from other survey data; and
- keep a log of the survey ID of all paper-based forms and where they are located; if documents are transported from one survey site or office to another, the log should indicate the survey ID of all forms in the shipment.

To secure electronic data:

- use access controls (password protection) on devices such as computers and tablets. Strong passwords have at least eight characters, including at least one lower case letter, one upper case letter, one digit and one symbol; a password should be required every time an electronic device is turned on;
- encrypt devices or the data stored on them;
- restrict access to the server to those who need it, and limit user permission to prevent unintended and undocumented changes to the data (e.g. view-only rights);
- control internet access and installment of new software to minimize the risk of malware installation;
- avoid storing computer data files with personal information on portable devices (laptops, tablets or flash drives) that can physically be lost;
- minimize transfer and use of portable devices (e.g. flash drives) to prevent physical loss of data and transfer of viruses or malware;
- ensure that portable devices that contain data are kept by field staff at all times when in the field, and are shut down when not in use;
- purge computers and other data-capture devices of data after the last interview of each day by uploading the collected interviews to a main, secure database, to minimize both the amount of data carried on the handheld device and the number of records lost or compromised if the handheld device is lost or stolen; and
- use an external source to back up data.

To secure electronic data during transmission:

- do not transmit any data over an unsecured Wi-Fi or internet connection;
- use a virtual private network (VPN) for connecting between sites and the server;
- avoid using email to transfer electronic files; if files must be emailed, they should be password protected or encrypted; and
- schedule regular antivirus software updates, software patch installs and virus scans.

When a computer used for data collection is taken out of service, any hard drives that may have once contained data must be wiped clean and reformatted before being used for another purpose. Software that repeatedly overwrite data on a disk with random patterns have been developed for this purpose and should be used. Other removable computer hardware (e.g. compact disks or DVDs used to store data backups) that are no longer needed should be destroyed and not used for another purpose. Upon conclusion of the survey, all documents containing sensitive information must be shredded before disposal. Similarly, all electronic files with sensitive information that are not part of the master dataset should be wiped clean and overwritten.

Breaches in the data security procedures should be reported immediately to survey staff leadership and investigated by staff to assess the causes and implement remedies. Confirmed breaches resulting in the release of sensitive information should be reported to the principal investigators, who must report breaches and any remedial action taken to all institutional review boards.

14. Staff selection

This chapter discusses assembling field staff for biobehavioural surveys (BBS), including their number, roles, and how survey activities are distributed between them. The roles and responsibilities of non-field-based staff (e.g. principal investigator and statistician) are not discussed here because these positions have usually been developed and filled before the decision is made to conduct a BBS.

A-14.1 Considerations for field staff selection

Staffing decisions are influenced by the survey's goals, objectives and activities. It is important to consider the number of field staff needed, their level of experience, their personal characteristics, and any accreditation they may need to collect and test biomarker specimens.

Number of field staff

The number of field staff needed at a survey site may depend on various factors, including the:

- complexity of the survey process (because a survey with more steps may require additional staff) and the sampling method used; for example, respondent-driven sampling (RDS) versus time-location sampling (TLS):
  - RDS is a passive sampling method, meaning that investigators have little control over how many potential participants to expect on any given day;
  - in TLS, survey staff spend a lot of time traveling to and setting up the survey site at the various locations;
- sample size (because the larger the sample size, the more interviewers will be needed to conduct the survey in the predetermined time frame);
- data-collection method – for example, audio computer-assisted self-interview (ACASI) or computer-assisted personal interview (CAPI) – and whether data are collected on paper or electronically; Paper-based data collection requires additional staff to double enter and clean the data;
- inclusion of a second-visit activity (e.g. follow-up questionnaire, provision of biological test results);
- security and transportation needs (which may necessitate security guards or additional staff, such as drivers for field staff); and
- method and type of biomarker specimen collection and processing (e.g. testing on-site or off-site).

Experience of field staff

The level of experience required from each staff member depends on the demands of their position. At a minimum, field staff should have some experience with the target population, and they should also have:

- previous experience working on other health surveys;
- knowledge of the topics and issues related to the survey;
- experience using computers or tablets if data are collected electronically; and
- experience in specimen testing and counselling if on-site testing is part of the survey design.

Characteristics of field staff

The personalities of staff members should be considered. Field staff will be more successful if they reflect some of the following characteristics:

- friendly and able to build rapport; able to make participants feel welcome and comfortable;
- responsible and reliable (e.g. arrive on time when the survey site opens, and take appropriate action when unexpected situations arise);
- mature and professional (e.g. able to interact with others in a professional manner);
- nonjudgemental and non-discriminatory toward the target population;
- able to work in a team;
- able to speak the language(s) spoken by the target population;
- committed to working for the duration of the survey; and
- available to work nontraditional hours (if required).

Consideration for inclusion of target population members as staff

Investigators may consider filling positions with members of the target population. This approach should be determined based on whether the formative assessment indicates a widespread acceptance of having community members serve as staff, because it has implications for survey participants' sense of confidentiality and safety. On one hand, community members may build rapport with survey participants more easily. On the other hand, survey participants may be concerned that staff members who are part of the community and their social network could compromise their privacy.

A-14.2 Field staff roles and responsibilities: non-method specific

Investigators must define staff roles and build teams within the context of the survey's objectives, budget and activities. As such, investigators may need to merge positions or divide them to make the best use of available resources, time and effort. They should adapt descriptions of the positions as appropriate for the country in which it is being conducted, the sampling methodology, and the survey site layout. For example, nurses may be nurse counsellors and also return results. In some settings and with adequate training, nurses may also collect blood and perform rapid testing. Table A-14.1 lists staff positions that are needed for all BBS, regardless of the recruitment method.

Table A-14.1 General BBS roles and responsibilities

Roles	Responsibilities
Field staff coordinator	Supervise daily activities at site Provide support to staff Communicate on regular basis with principal investigator
Interviewer	Know the questionnaire and where sensitive issues may arise Build rapport with survey participants Obtain informed consent from eligible survey participants Conduct interview or oversee the interview process Record responses Ensure that information collected (including paper questionnaires and electronic data) is securely saved or stored after the interview
HIV counsellor/nurse or laboratory technician	Provide pre- and post-test HIV counselling based on local regulations Collect biological specimens Complete relevant forms and processes, and package specimens for testing or transport Conduct on-site rapid testing as specified in protocol
Data manager	Manage data entry and database development Oversee data editing and quality control
Data entry staff	Capture data and enter it into database
Data analyst	Analyse and interpret data
Laboratory staff	Test specimens Record test results Transfer results data to main survey database
Community liaison/mobilizer	Meet with community members Share survey progress with members Encourage participation in survey
Driver	Transport staff and supplies
Support staff	Protect field staff and participants Maintain survey site

A-14.3 Field staff roles and responsibilities: method specific

Some BBS methods require specific additional positions to be filled. Tables A-14.2 and A-14.3 list positions specific to TLS and RDS. Because TLS and RDS survey sites are not investigator-controlled environments, it is important for the survey investigators and field staff to meet and work together before survey implementation.

A-14.4 Considerations for staff payment

Funding for paying staff salaries should be accounted for in the budget as early as the survey-planning phase.

Survey coordinators should research pay scales and payment schedules of comparable survey projects in the respective country or region. Payments should be comparable to the salaries of similar local jobs, and decisions about salary rates, payment schedules and working hours should be made by someone who is knowledgeable about the local context.

Investigators should also consider roles for staff that are remunerated in-kind rather than paid. Local laws and employer policy will govern contractual arrangements, sick and annual leave, and termination of contract.

Table A-14.2 TLS-specific roles and activities

TLS specific	
Roles	Responsibilities
Enumerator or counter	Collect quantitative data used to determine inclusion of venue in survey Count venue attendees and direct recruiters to approach selected attendees
Recruiter	Approach venue attendees Explain survey to potential participant, describing: • the purpose • interview procedures • privacy protection • compensation

Table A-14.3 RDS-specific roles and activities

RDS specific	
Roles	Responsibilities
Receptionist	Greet people entering the survey site Check that people entering the survey site have a valid coupon Inform people how the coupon process works when they arrive without a coupon Give information about the survey over the phone or at the survey site Track and manage the flow of participants Inform participants of how long they may have to wait for an interview
Eligibility screener	Screen participants to see whether they fulfil the eligibility and inclusion criteria Ask specific questions about behaviours related to the target population to identify and disqualify imposters Initiate checklist form Establish a unique identifier to link a recruit's follow-up visit to initial visit Verify coupons
Coupon manager	Develop coupon-tracking system (if not already existing) Manage the coupon-tracking system Explain the survey to the participant Conduct and record consent Give instructions on peer recruitment Issue referral coupons Provide primary or secondary compensation, as warranted

Section A

15. Standard operating procedures

This chapter provides guidance for creating and using standard operating procedures (SOP) for a biobehavioural survey (BBS). An SOP is a set of written instructions for distinct survey activities. Developing and using SOP helps to ensure that staff uniformly conduct the survey as recommended by the investigators and as outlined in the protocol. SOP help to maintain quality and consistency, particularly when staff change.

Key terms

Standard operating procedure:	A set of written instructions for distinct survey activities.

An example SOP that can serve as a template is provided in Appendix I-30.

A-15.1 How to write a standard operating procedure

Style

SOP are written in concise, clear, simple language (active voice and present tense) using a step-by-step approach. Flow charts can help illustrate steps for staff to follow.

SOP preparation

Survey investigators determine which activities or processes warrant an SOP. An SOP should be written by individuals knowledgeable about the survey design and activities (i.e. principal investigator, survey coordinator, data or laboratory manager), and should contain enough detail for staff to carry out the procedure without supervision.

Each page of the SOP should have a header or footer, or both, with:

- a unique title; and
- an approval date, a version number, or both.

The first, or title, page should list:

- the staff concerned;
- a summary of content, with purpose and field application;
- any related SOP;
- the name of the author; and
- the name of the person authorizing the SOP (optional).

SOP review and approval

The draft SOP should be reviewed by other individuals who have experience with the activity being described. Survey implementers should field-test complicated SOP, and investigators should approve the final version of the SOP.

SOP revisions

SOP need to remain current in order to be useful. When procedures change, SOP should be updated and reapproved. An SOP change log can be used to track these changes and the versions approved. SOP should be reviewed periodically during the survey to ensure that procedures are still appropriate. The survey coordinator is responsible for maintaining and updating SOP as needed. After an update, the survey coordinator should provide the revised SOP and associated training to all relevant staff, and ensure that staff are following the revised SOP.

A-15.2 BBS activities that require an SOP

Many survey activities and procedures can benefit from a written SOP; for example:

- selecting and training field staff (e.g. training in sampling; see Table A-15.1 for examples);
- recruiting staff and eligibility screening;
- obtaining consent;

- ensuring confidentiality;
- standardizing field site operations (participant flow, hours of operation and appointment system);
- collecting data;
- monitoring data collection;
- developing survey identification numbers (ID);
- providing test results to participants;
- providing HIV counselling;
- providing or referring for treatment;
- treating sexually transmitted infections (STI) or vaccinating for hepatitis B;
- reimbursing funds;
- ensuring safety of participants and staff;
- developing personal evacuation plans;
- dealing with adverse events;
- dealing with repeat participants or imposters;
- dealing with survivors of violence;
- undertaking laboratory activities:
  - specimen collection and long-term storage;
  - specimen transfer to the laboratory;
  - algorithms for specimen testing ;
  - laboratory quality-control procedures;
  - communication of laboratory results to survey site;
- managing data (including data entry and data cleaning); and
- ending data collection or closing survey site.

Appendix I-27 includes an example SOP on data-management for RDS surveys using Questionnaire Development System (QDS).

A-15.3 Sampling-specific SOP

Some sampling methods include activities that require specific SOP (see Table A-15.1).

A-15.4 SOP for checklists, logbooks labels and forms

Various types of survey-specific instruments (e.g. checklists, forms, logbooks and labels) should have an SOP that provides detailed instructions for making use of the instrument. Some SOPsinclude checklists for staff to use when conducting a BBS; such SOP help staff to identify important survey tasks during BBS implementation. Table A-15.2 gives examples of these instruments and the corresponding SOP.

Table A-15.1 Activities associated with sampling methods that require standard operating procedures

Sampling method	
Conventional cluster sampling, time-location sampling	Respondent-driven sampling
Sampling frame preparation (list of venues) and observation of venues	Seed selection
Enumeration, cluster information collection	Coupon distribution and management
Participant selection, listing of people	Peer recruitment training
Participant recruitment	Reimbursement management
Logistics of prepping for each day of collection	Managing multiple survey sites (e.g. multiple sets of staff or one set that moves among sites)
	Welcoming to survey site. Waiting room activities
	Visiting procedure (e.g. what happens at first and second visits)

Table A-15.2 Survey-specific instruments that require an SOP

Instrument	Examples	Corresponding SOP
Checklist	• Survey procedure checklist • Field supervision monitoring checklist	• Flow planning of survey procedure
Logbook	• Survey ID log • Appointment log • Specimen collection log • Specimen shipping log • Specimen registration log (at laboratory) • HIV test results log • Log for other biological tests • Supply lists	• Specimen collection and transfer
Label	• Unique ID • Specimen label	• Specimen collection and transfer
Form	• Participant tracking form • Cluster information sheet	• Flow planning of survey procedure • Sampling

16. Training and training methods

This chapter provides guidance on staff training and the information that should be covered in the training materials. Training should familiarize staff with the goals and purpose of biobehavioural surveys (BBS) and provide background information on HIV and treatment.

At the end of the training sessions, staff should be aware of:

- steps in the survey implementation process and related tasks;
- roles and responsibilities of each staff member involved in the survey;
- sampling methodology;
- laboratory procedures;
- data collection and management;
- ethics and confidentiality; and
- safety and security procedures.

A-16.1 Training

Training should also increase staff comfort in interacting with participants and discussing sensitive topics in the BBS questionnaire. Trainers should instruct on all BBS standard operating procedures (SOP), and should make these and all other training materials (e.g. copies of data instruments, protocols and role-play scenarios) available to staff members to serve as a reference. The cost, required hours and location of training should be estimated during the design phase and included in the budget (see Appendix I-2).

In addition to survey procedures and tasks, roles and responsibilities, and SOPs, other important topics include staff behaviour, how to conduct interviews, specimen handling, and safety and security.

Professional conduct

Staff should conduct survey activities in a manner that adheres to ethical research principles, by respecting and protecting the privacy, confidentiality and autonomy of participants; also they should sign a confidentiality agreement (see example in Appendix I-29). Survey staff should remember that survey participation is voluntary, and should be trained in obtaining informed consent and in the consequences of breaching participant confidentiality. They should also be trained in professional behaviour and demeanor towards fellow staff and survey participants. Staff should not "date", have sexual intercourse with or share drugs with participants. Staff should also complete cultural competency and sensitivity training in order to learn how to work with highly stigmatized groups, and should be trained in positive body language and good interpersonal skills.

Interviewing

The attitude of interviewers can influence participants' answers in face-to-face interviews, especially in interviews that ask about illegal or stigmatized behaviours. To encourage honest responses, interviewers must be thoroughly trained in open and nonjudgemental questioning techniques and accurate recording of responses.

The amount and type of training required will vary depending on who is conducting the interview. For example, if the peers of those in the participant group are selected as interviewers, they may be less likely than professional staff to appear judgemental. Conversely, they may be biased towards recording or coding responses in a way that reflects their own opinions or behaviour. If the questionnaire is electronic and self-administered, staff should be trained in how to use the equipment, how to assist a participant to complete the questionnaire, and how to troubleshoot problems with computer equipment and the electronic instrument interface.

Biological specimen collection, handling and transportation

Detailed information about how to collect, handle and transport biological specimens to the laboratory should be included in SOP. Survey staff responsible for biological specimen procedures should be thoroughly trained in these procedures. This training should include a practice run ("dry run") to detect and address issues that may arise.

Safety and security

The safety and security of the survey team and participants should be a high priority for investigators. Survey staff should be trained in safety procedures and should practise how to respond to potential security threats, including requests for information or entry to a survey site by police or other authorities. Appendix I-9 has examples of security SOP.

A-16.2 Training methods

An interactive practicum-style setting helps with learning and allows trainees to rehearse survey procedures and tasks. Role-playing and simulating (i.e. practising) survey events from beginning to end are two methods for training staff.

Role play

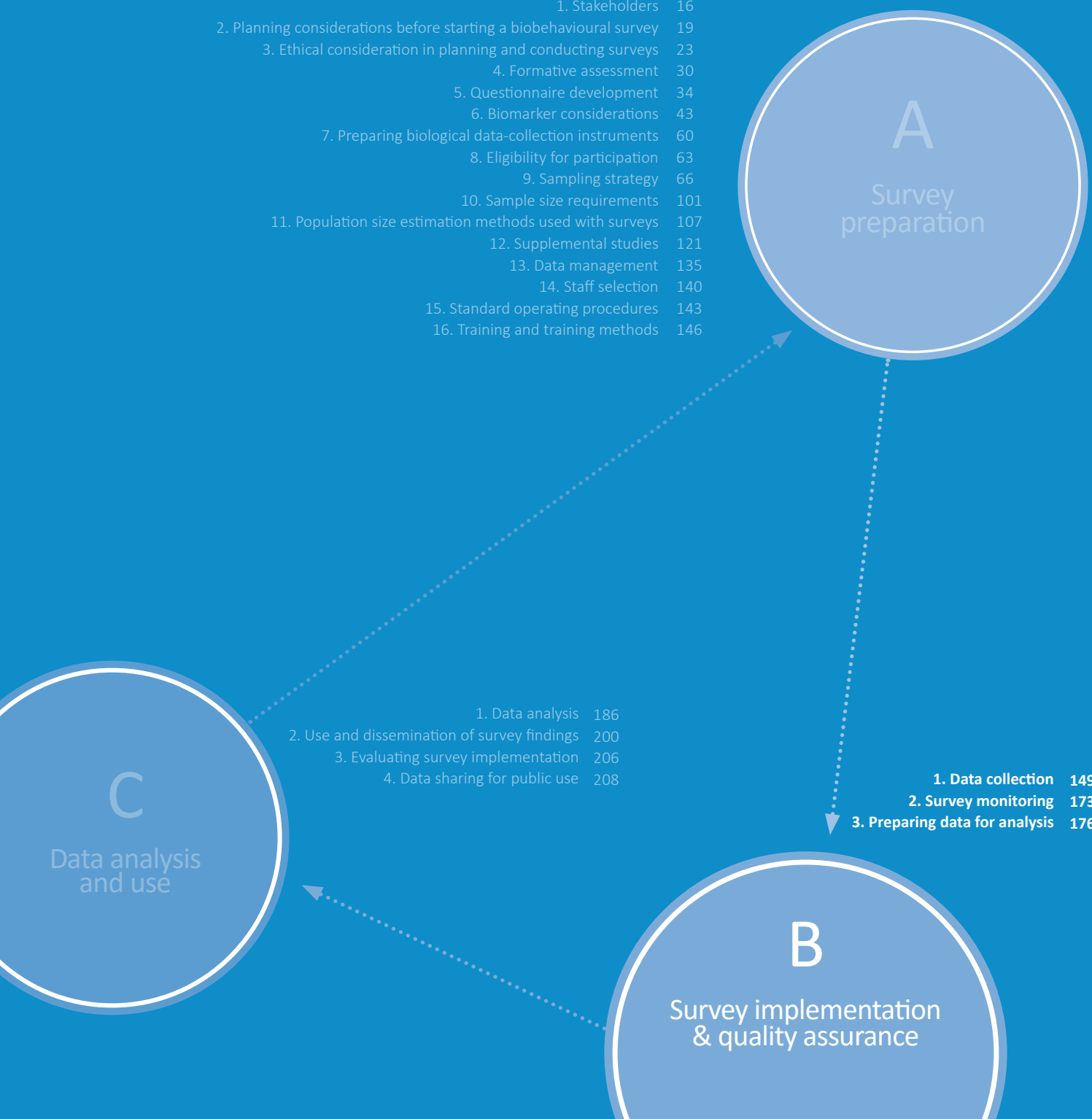
Role playing allows staff to practise scenarios and demonstrate their knowledge. Staff should participate in role-play exercises after completing training sessions on qualitative and quantitative interviews, consent administration, eligibility screening, reception of recruits (for RDS), counselling and testing, and returning results (if applicable).

Survey procedure simulation

Once staff have been trained in their respective tasks, a practice session of survey activities, using a mock survey in a target population similar to the intended survey population, is a good exercise for understanding the survey event process. The practice should follow the prescribed flow of events and be as realistic as possible. After completing the exercise, the trainer should discuss staff knowledge of roles and performance of activities.

B. Survey implementation and quality assurance

1. Stakeholders 16
2. Planning considerations before starting a biobehavioural survey 19
3. Ethical consideration in planning and conducting surveys 23
4. Formative assessment 30
5. Questionnaire development 34
6. Biomarker considerations 43
7. Preparing biological data-collection instruments 60
8. Eligibility for participation 63
9. Sampling strategy 66
10. Sample size requirements 101
11. Population size estimation methods used with surveys 107
12. Supplemental studies 121
13. Data management 135
14. Staff selection 140
15. Standard operating procedures 143
16. Training and training methods 146



1. Data analysis 186
2. Use and dissemination of survey findings 200
3. Evaluating survey implementation 206
4. Data sharing for public use 208

1. Data collection 149
2. Survey monitoring 173
3. Preparing data for analysis 176

1. Data collection

This chapter discusses the implementation and monitoring of nonprobability surveys. As mentioned in Section A (Survey Preparation), nonprobability sampling is the method of choice when more rigorous sampling designs are not possible due to cost, complexity or relative inaccessibility of the target population.

B-1.1 Surveys using nonprobability sampling methods

Key terms

Community-based organization: An organization largely comprised of members from a given population (e.g. men who have sex with men), often working in advocacy or providing services.

Community liaison: A person acting as a link to a target population. May be a representative of a population.

Drop-in centre: A safe venue that provides health, educational or counselling services to a particular population (e.g. people who inject drugs).

Gatekeeper: A person who controls access to people or a place.

Because the probability of selection for a given participant is unknown in nonprobability surveys, sampling bias poses the largest risk to the representativeness of such surveys. For example, if a survey is designed to sample sex workers (SW) at clinics treating sexually transmitted infections (STIs), a substantially higher HIV and STI prevalence would be expected among this sample compared to a survey that samples SW at venues where SW solicit clients. Although such nonprobability samples do not allow inference to the larger target population, survey investigators should still strive for a varied sample to potentially reduce selection bias (e.g. by sampling from as many STI clinics as possible, and from other venues). Before planning a nonprobability sampling survey, investigators should consider the limits of nonrepresentative survey results and the allocation of time and money to the survey.

In most cases, investigators choose the sampling method before developing the protocol. Sometimes, however, a probability-based survey may fail to produce a representative sample. Possible reasons include differential recruitment, refusal of venues to provide access, and sparse social networks. The survey then becomes, in effect, a nonprobability-based survey and the data should be analysed as such. Documenting survey procedures and results that prompt the decision to treat the data as nonrepresentative helps investigators to adjust for these conditions and better address biases. It also helps readers to assess the validity of results.

B-1.1.1 Convenience sampling

Convenience sampling is perhaps the simplest nonprobability sampling method because individuals are recruited for participation based on the ease of finding them. For example, a survey may consecutively enrol people who inject drugs (PWID) who frequent a community-based organization or drop-in centre. Convenience sampling may take many forms, all of which generally facilitate sampling for the investigator. Although selecting the sample through multiple means (e.g. referral, at an STI clinic or brothel, and through advertisements at bars or clubs) may produce more diversity in the sample, it cannot be regarded as probability-based.

Take all approach

A very simple version of convenience sampling is the “take all” approach, in which all eligible persons identified are offered enrolment. Examples of this approach include sampling all eligible SW in a brothel, all transgender women (TG) in a bar or all PWID frequenting a drop-in centre. If the number of venues is too large, the aim should be to include as many as possible with

as much variety as possible to help diversify the sample. First, all known sampling locations should be listed, then a feasible number selected, and all eligible persons encountered at these venues enrolled.

B-1.1.2 Purposive sampling

Sometimes a survey targets a particular subset of a population such as men who have sex with men (MSM) who are also married to women, or HIV-infected members of the target population. The implementation of a survey using purposive sampling depends on the desired sample characteristics. For example, a survey targeting MSM who are married to women may require the use of peer referral, whereas a survey targeting HIV-infected members of the target population may sample people at an antiretroviral therapy (ART) clinic. Purposive sampling may also be used to select particular attendees at an ART clinic for participation. For example, if investigators find that most people come from only one part of the city, they may preferentially recruit those from other parts so that the sample is more reflective of the whole population.

B-1.1.3 Quota sampling

Quota sampling is similar to purposive sampling, except that it aims to sample a given number of people with specific characteristics. This form of sampling may be used to ensure inclusion of a sufficient number of people with characteristics of interest, or to ensure that the sample is more reflective of the whole population. For instance, investigators may choose to sample 100 PWID who know their HIV status and 50 PWID who do not. PWID of known HIV status may be found by collaborating with care providers or social gatekeepers. PWID of unknown status may be sampled (selectively) through chain-referral sampling (described below) or at drop-in centres.

Other examples include having quotas for the number of street-based and brothel-based SW, or for SW born domestically and internationally.

B-1.1.4 Chain-referral sampling or snowball sampling

Snowball sampling may be a good option when target members know one another. Snowball sampling begins by purposively sampling some initial participants who are then asked to help recruit others. The purposively sampled participants may share with the staff how and where to locate their peers; alternatively, the staff give the participants vouchers and information about the survey and ask them to either recruit or spread awareness of the survey to their peers. Additional details on using vouchers (coupons) for sampling are available in Section A-9.7 on respondent-driven sampling (RDS).

Snowball sampling differs from RDS in that, in snowball sampling:

- there is no limit on the number of coupons issued per person;
- the recruiter–recruit relationship does not need to be tracked;
- recruiter and recruit do not need to know one another; and
- survey staff may make initial contact with potential participants.

B-1.1.5 Recruitment strategies

Potential survey participants can be recruited from a target population in various ways, including at commercial venues or public spaces, and through service providers, community-based organizations or the internet, as discussed below.

Commercial venues and public spaces

Sampling at venues and other locations may be preferred if many or most members of the target population frequent these locations. Venues may include bars, clubs, parks, hotels, brothels and selected streets. Usually, the more kinds of locations sampled, the less biased the sample. Survey investigators may need to ask venue owners for permission to sample at the venue. See the section on time-location sampling (Section A-9.6) for discussion of the practicalities of sampling at venues.

Service providers and community-based organizations

Service providers can help identify target population members. Service providers may include individuals working for STI clinics, community-based organizations that provide services, drop-in centres, nongovernmental organizations (NGOs) or other groups that distribute prevention materials (e.g. condoms, lubricants, needles and syringes). Survey investigators may offer enrolment either to a portion or to all service recipients or community-based organization members or attendees. If feasible, potential participants should be sampled directly at the service delivery site or office of the community-based organization to facilitate the speed of recruitment. If direct sampling is not possible, service providers and staff of community-based organizations can distribute vouchers or flyers inviting individuals to participate in the survey at another location. Survey investigators may also enlist the help of outreach workers who know where to find or contact members of the target population; these workers may enjoy more trust and have better rapport with the target population than do survey staff.

Internet-based surveys

Websites of interest to a target population (e.g. gay-friendly websites) can also be used for recruitment. Investigators may place a link on a website that redirects

the user to the survey webpage. Before the actual self-administered interview, participants undergo self-administered eligibility screening and consent. Some internet surveys also ask participants to refer their peers; for example, by providing an email address that is used to automatically send an invitation for survey participation.

Participants' internet protocol (IP) addresses should be considered as personal identifiers; survey investigators must ensure that the addresses are protected if they are collected or stored, or that participants can remain anonymous.

B-1.1.6 Preparation for sampling

Each sampling method and recruitment strategy has unique operational requirements, but the considerations below apply to all methods in some way:

- contact gatekeepers and describe the survey objectives and confidentiality.
- set expectations for the survey activities with the gatekeeper. Explain (orally or in writing, as appropriate) the requirements for listing individuals at the location; the physical space needed for interviewing, collecting specimens or providing test results; the time needed for each interview; and any testing procedures.
- identify a private area appropriate for conducting interviews and collecting biological specimens.
- train survey staff on protections for participants, including ensuring that participants understand and sign confidentiality agreements.
- schedule participant sampling or interviews.

B-1.1.7 Enrolment and data and specimen collection

Every effort should be made to ensure data quality and minimize errors. Measures to ensure security for staff and participants should also be considered. Most surveys, including those that employ nonprobability sampling and those with more rigorous sampling designs (see Section B-1.2) use the same procedures (e.g. consent, confidentiality measures, interviewing, specimen collection and testing). Investigators should follow the survey protocol and flow chart of survey activities. Enrolment may take place at the location of sampling (e.g. at a venue or in a van, a room in a clinic or a park bench). Other surveys may refer or transport potential survey participants to a fixed survey office with better security. Sampling and enrolment may need to occur in the evening or late at night when the target population may be more accessible.

Depending on where data and specimen collection take place, survey staff should be prepared to administer surveys and collect and test specimens in the field; to do so, they may need to carry backpacks, cooler boxes, flashlights, clipboards and tablet computers with them. If feasible and affordable, a van or bus can provide the mobility needed for sampling at venues, facilitate data and specimen collection, and increase security and privacy. Regardless of the type of data-collection activity, relevant staff require certain items. These include copies of the letter of support from the ministry of health, and institutional review board (IRB) approvals; sufficient paper copies of the data collection forms; sufficient and appropriate supplies for specimen collection and transport; information or brochures about HIV and STI prevention, care, treatment and services; and reimbursements for participants (e.g. money in a cash box, condoms, lubricants and clean injection kits).

B-1.1.8 Monitoring of survey progress

Adherence to standard operating procedures (SOPs) is equally important for surveys that use nonprobability sampling and those that use probability sampling. Documenting and monitoring survey progress and sample composition is required to gauge both the extent of various biases and data quality. SOPs should be developed and changes documented. As far as possible, investigators should document how sampling was conducted, including the sampling universe (i.e. the number and location of sampling sites, and the proportion of each included in the survey). They should also record data that inform the potential denominator (i.e. the probable size of the target population accessible through the planned sampling design).

During the sampling process, staff should document the number of people screened for eligibility, the number who are eligible, the number who consent, and the reasons for refusal. They should also document the number of participants who agree to specimen collection and the number of people who receive their results, as well as reasons for refusal. If vouchers or coupons are used, the numbers issued and redeemed should be tracked.

In addition to documentation of the actual sampling procedures, attention to data and specimen quality is important. Staff should note the potential for non-sampling error, such as reporting bias in the interview process or testing quality in a mobile or outdoor environment.

B-1.2 Conventional cluster sampling

This section provides guidance on using conventional cluster sampling (CCS). Procedures explained here follow the activities described in Chapter A-9 about creating the sampling frame and selecting clusters.

This section also describes preparations for participant sampling, special confidentiality and anonymity procedures, field staff arrival and set-up, methods for sampling individuals from a list, recruitment techniques and other survey procedures.

In CCS, the target population is associated with clusters in a "fixed" manner. That is, individuals are associated with only one cluster location, the same individuals can reliably be found at the same cluster regardless of when it is visited, and a complete list of individuals who are associated with a cluster can be obtained or created. Appendix I-20 provides additional guidance on mapping cluster locations and creating lists of potential participants.

Examples of survey situations in which CCS might be appropriate include SW working in brothels or homes, detainees in prison, military personnel in barracks, and PWID in treatment facilities.

Key terms

Actual measure of size:	The actual number of individuals who meet the approach criteria during a sampling event.
Approach criteria:	Criteria for who, among all the venue attendees, staff should approach for participation in the survey.
Community liaison:	A person acting as a link to a target population. May be a representative of a population.
Counter:	Device used by enumerator to count people who cross into the recruitment area; or a person who conducts the enumeration.
Enumeration:	Counting all people in a recruitment area.
Gatekeeper:	A person who controls access to people or a place.
Intercept:	Approaching or engaging a potential participant with the expectation of screening the person for survey eligibility.
Recruitment area:	Location where recruiters approach potential participants to invite them to participate in the survey.
Sampling event:	The period in which field staff are in a venue recruiting and interviewing respondents.

B-1.2.1 Preparation for sampling

The steps listed below take place during the survey planning phase after the cluster locations have been selected from the sampling frame. These steps relate to the logistics of participant sampling at each selected cluster location.

- contact a gatekeeper at each selected cluster location and describe the survey objectives.
- visit each selected cluster location to build rapport, understand potential field issues, answer questions and address concerns.
- observe the physical layout of the location and \ work with the gatekeeper to identify a private area appropriate for conducting interviews and collecting biological specimens.
- discuss confidentiality with key people in the selected clusters.
- set expectations for the survey activities with the gatekeeper. Explain (orally or in writing, as appropriate) the requirements for listing individuals at the cluster location; the physical space needed for interviewing, collecting specimens or providing test results; the time needed for each interview; and any testing procedures.
- request a complete list of eligible individuals who are connected to the cluster location. If the plan is to recruit a stratified sample of participants (e.g. by gender or age group), request a list that includes the information needed to stratify the sample.
- schedule participant sampling date.

Special confidentiality considerations

Several special confidentiality concerns should be considered with CCS. First, gatekeepers do not have the right to see information collected from survey participants. Second, the information on the list of eligible individuals should be specific enough that any one individual can be identified among all individuals in the cluster. To protect the confidentiality of individuals on the list, investigators must follow these procedures:

- after sampling is complete (i.e. after all selected individuals in the cluster have been asked to participate), destroy the list.
- gatekeepers may try to claim the right to see the information collected from survey participants, including HIV test results. Human subjects protections dictate that this information not be shared. If a

gatekeeper insists on access to the survey data in exchange for sharing a list of eligible individuals, do not include this cluster location in the survey.
- train survey staff on protections for participants and ensure they understand and sign confidentiality agreements.

The following subsections describe set-up, screening and interviewing, and closeout of sampling events via specific steps that should be taken after the team arrives at the selected cluster location.

Preparing for sampling events

Step 1: Meet with gatekeeper

Ensure that the gatekeeper knows the purpose of the survey and the activities that will be taking place. If a community liaison is part of the team, ensure that person is part of this meeting.

Step 2: Obtain or create a list of individuals in the cluster

If a list of individuals associated with the cluster location is not available, ask the gatekeeper to create one. The investigator or other responsible supervisor can support the gatekeeper and other key informants at the cluster location to create the list.

Characteristics of a good list of individuals in a cluster include:

- each individual associated with the cluster is on the list.
- no individual is listed more than once.
- only eligible individuals are listed.[1]
- each individual can be identified based on the information on the list.
- if stratification is being used at this stage, the elements needed to stratify the sample should be included on the list (e.g. if stratifying by age, the list should specify the age of each potential participant).
- the list includes space to specify which selected individuals are absent or unavailable on the day the survey takes place.

Setting up at sampling events

Step 1: Set up the space for conducting interviews and HIV testing

The physical space for conducting interviews and HIV testing and counselling should be out of view and earshot of others. Control the space to prevent onlookers or others from invading the privacy of participants.

Random selection from a list of participants

Participant selection should be a quick and simple process. The objective is to give an equal chance of selection to all of the eligible individuals on the list; that is, to randomly select people within a cluster. These are two of the many ways to randomly select individuals:

- *Systematic random sampling:* Divide the total number of people on the list (N) by the number needed to participate (n). Select each n/Nth person on the list. **Note:** It is best practice to roll a dice or to use some other random-number generator to identify a starting point on the list. For example, if a rolled dice shows a "4", start counting every n/Nth individual with the fourth person on the list.
- *Random sampling:* Create N (the total number of people on the list) pieces of paper with sequential numbers from 1 to N and put them in a container. Without looking, select papers until the number needed to participate (n) is selected. Match the number on the paper to the ordered number of people on the list. For example, if the number "5" is chosen, recruit the fifth person on the list. This method is more appropriate when a smaller number of eligible participants is needed.

Step 2: Select individuals for participation from the list

Select the required number of individuals (including those who are eligible but not present at the time of selection) according to the procedures laid out in the protocol (see Chapter A-9). To avoid bias, sampling is done solely by the supervisor, not by the community liaison.

Note: After the initial sample of individuals has been approached, it may be necessary to re-sample individuals if the number of refusals is higher than expected – this step is only necessary when the number of refusals prevents the team from enrolling the target number of participants.

[1] If eligibility includes behaviours, the method for determining eligibility for individuals who have a fixed association with a site should be described in the protocol and operations manual. If it is not possible to obtain a list of only those individuals who are eligible, the survey instrument should include an eligibility screener to determine the individuals who may be interviewed.

Example of recruitment talking points

Thank you for your interest in our survey. I would like to tell you about the survey:

- We will describe the survey procedures in detail before we collect any information from you.
- Being part of the survey is voluntary and confidential.
- The survey includes an interview and HIV testing. The interview will take up to 60 minutes to complete.
- We will ask you questions about your background, and about behaviours surrounding sex and substances you may use.
- We would like to take a blood sample for an HIV test.
- There are minimal risks to being part of the survey.
- By being part of the survey, you will help us to plan better HIV services for your community.
- You will receive condoms, lubricants and referrals to services, if you wish.
- We will answer any questions you have about the survey before you start.

Sampling individuals from a list

Investigators should select from the complete list of individuals associated with the cluster location. There may be situations in which a number of eligible individuals on the list are not available to participate when the survey is conducted. In this case, do the following:

- Schedule an alternative time to survey those who are not available. This alternative time should be as close as possible to the initial survey date. Staff should make at least three attempts to complete the survey with those who are unavailable before seeking a replacement participant.
- Document selected participants who are not available to participate as "nonresponse".
- Document the "measure of size" or the total number of eligible individuals associated with the cluster.

Recruiting, screening, consenting, interviewing, counselling and testing

Step 1: Approach and recruit selected individuals
This step may be combined with or designed to complement Step 5 in below.

The supervisor and, if appropriate, the community liaison speak with each person selected from the list. Rather than talking to all potential participants as a group, aim for one-on-one conversations to ensure an individual is free to decide about participation without pressure from others. Build rapport with the potential participant; discuss the survey, including its purpose and benefits; and explain the random selection process and confidentiality protections. During this step, it may be helpful to use recruitment "talking points" to briefly describe survey procedures, risks and benefits, and other elements of the consent process.

The field team may have fewer interviewers than the number of selected participants in a cluster location. For example, the team may have two interviewers but need to interview five individuals from a cluster. In this case, the team should first approach two of the potential participants to conduct the interviews, and when an interviewer is free, then approach the third participant. This technique helps to ensure that participants do not have to wait for an interview to begin.

If selected potential participants are not present, schedule alternative times to conduct the interviews with them.

Step 2: Escort individual to interview area
Escort potential participants to the physical location where interviews and HIV testing will take place.

Step 3: Assign a survey identification number
The specific procedures for assigning survey identification numbers (ID) and deciding which staff member is responsible for this step be included in an SOP. Survey ID assignment is commonly done by one staff member to avoid duplication of IDs. If more than one person is assigning survey ID, allocate different survey ID numbers to each relevant staff member.

Step 4: Screen for eligibility
Ensure the selected potential participant is eligible by screening the person using a standardized tool. Document eligibility for each individual approached. If someone is ineligible, inform the person and offer thanks for their time.

Note: Supervisors should monitor the number of ineligible individuals logged by each interviewer for quality improvement purposes. Interviewers who log an unexpectedly high or low number of ineligible persons may require retraining or mentoring to ensure proper screening procedures.

Step 5: Obtain informed consent
Obtain informed consent according to procedures described in the SOP and document the outcome of the consent process (see Chapter A-3 for more information).

Step 6: Interview the participant
Administer a face-to-face or self-administered interview using an electronic questionnaire. If both methods are offered, document the interview mode used.

Step 7: Conduct counselling and testing
The interviewer escorts the participant to the area designated for counselling, and specimen collection and testing. Provide pre-test counselling and collect biological specimens. Conduct post-test counselling according to the protocol and SOP. Ensure that the result of each test in the testing algorithm is recorded. Return test results if conducting rapid testing, or schedule a date to return test results.

Step 8: Provide referrals as appropriate
Participants newly found to be HIV positive and those who are not in care for HIV infection must be referred for medical care and treatment as described in the SOP. All individuals testing negative for HIV should be given information about where they can be tested in the future.

Participants in need of other health or social services should also receive referrals to the appropriate service providers in the community.

Step 9: Provide treatment for treatable STIs
Conduct this step if it is part of the protocol, according to the SOP for STI treatment.

Step 10: Provide prevention materials
Provide participants with HIV-infection prevention materials and supplies (e.g. informational pamphlets, condoms, lubricant and clean injection kit).

Step 11: Provide compensation as appropriate
Provide participants with compensation if this is part of the protocol and SOP.

Closeout

Debriefing
At the end of data collection at the cluster location, a debriefing for field staff is useful in order to discuss any issues or problems that occurred. This debriefing can be either a meeting convened with all field staff or a series of one-on-one conversations. Possible questions for discussion include:

- In general, how did data collection go at the selected cluster?
- Were there any cluster-related issues that affected survey activities?
- Were there specific concerns related to selection of the participants, consent process, and so on?
- Were there any barriers to recruitment? What strategies were successful in overcoming barriers?
- Were there any unusual or adverse events (e.g. a participant ended the survey early, or a participant initially consented to an HIV test but changed their mind later)?
- Were there any problems with the electronic devices?
- Were there any possible errors within the survey data?
- Were there any problems with specimen collection or test kits?

If necessary, the field supervisor can record this information for future reference. The field supervisor should notify the principal investigator or other staff if an SOP or the protocol requires amending.

Field notes
The field supervisor should document any problems, barriers or challenges at the cluster location where the data were collected. This information is important when investigators are reviewing recruitment statistics and can help to adjust field operations in future clusters, if necessary.

Collect and review forms and logs
At the end of the data-collection activities at a selected cluster, the field supervisor should collect all materials used by field staff, review them for accuracy, make corrections as necessary, and tabulate recruitment outcomes (e.g. number of individuals selected, number approached, number accepting participation, number of refusals and reasons for refusal).

B-1.3 Implementation of time-location sampling

This section describes the implementation of surveys using time-location sampling (TLS), including the core activities of preparation for the sampling event: event setup, enumeration, sampling, enrolment and other survey procedures. As described in Section A-9.6, TLS is used to survey populations who come and go freely from physical venues (locations) but may not visit the venues regularly. Although it would be possible to create a list of all people at a venue, the list would vary depending on the day and time the location is visited by the survey team.

In TLS, investigators create a sampling frame comprising all possible venues and times that the target population may visit. Investigators then sample from a list of venue–day–time combinations to obtain a representative sample of the people attending venues. Survey staff then visit the venues to collect data during "sampling events". Individuals present at a venue during a sampling event are eligible to be sampled in the survey. Most of the people at the venue during the sampling event should be part of the survey's target population.

TLS may be appropriate for surveying SW waiting for clients at bars, nightclubs, street locations or hotels; MSM or TG attending bars, clubs or bathhouses; and PWID attending drug-buying or drug-using venues or other similar open spaces.

B-1.3.1 Survey staff

For each sampling event, the survey team should include a combination of the following staff positions: supervisor, counter, recruiter, interviewer, counsellor and laboratory technician. Ways in which these positions can be combined are described below. For details regarding the staff involved in a TLS survey, see Chapter A-14.

B-1.3.2 Preparation for the sampling event

The following steps take place **after** the venue day–times have been selected (see Section A-9.6). The steps relate to the logistics of participant sampling and recruitment at each site.

Step 1: Meet with the gatekeeper of each selected venue and, if appropriate, the community liaison
Gatekeepers should be engaged whenever possible to facilitate access to the target population and survey planning. As survey activities may interfere with the normal routine of a venue, a gatekeeper's acceptance of or support for the survey may be needed. Gatekeepers may include bar or brothel owners, pimps, nongovernmental organization (NGO) staff, or drug dealers. In some cases, there may be not be an obvious gatekeeper.

Together with the community liaison, the field team should meet with the venue gatekeeper to inform them about the objectives of the survey and the sampling and survey procedures. A community liaison is a person acting as a representative of or a link to a population to facilitate survey planning, implementation, interpretation and dissemination of results. For a park or street-corner venue with no obvious gatekeeper, the field team should work with the community liaison to identify any key people to meet.

Because the gatekeeper may differ each time the field team visits a venue, be prepared at every meeting at the venue to meet the gatekeeper present and do the following:

- Explain the goal of the survey, including that it will benefit the target population and that their privacy and confidentiality will be protected.
- Set expectations for the survey activities.
- Explain requirements for enumerating and intercepting individuals at the venue, including that data collection may occur inside the venue, and the need for privacy for interviewing and testing.
- Emphasize that the survey team will do its work with minimal possible interference with "normal business" at the venue. Remind the gatekeeper that all survey data are confidential and cannot be shared with the gatekeeper or venue staff, if they exist.

Step 2: Prepare materials for the sampling event
Ensure that all supplies are taken to the sampling event. Below is a non-exhaustive list of important items to be taken to every sampling event:

- letters of support from the ministry of health and institutional review board (IRB) approvals: take at least two copies of each document to show potential participants, gatekeepers, or authorities, if needed.
- information and brochures about prevention, care, treatment and services for HIV and sexually transmitted infections (STIs); compensation for participants (e.g. money in a cash box), and condoms, lubricants, clean injection kits.
- copy of the venue observation form for reference (see Appendix I-20).
- data-collection instruments and forms: approach forms, consent forms, questionnaires, test-result forms,

referral forms, sampling event forms, incident report forms, and counting device. Bring paper backups of all electronic documents.
- electronic hardware (if applicable): tablets, cell phones, internet modems.
- laboratory supplies: specimen-collection devices (e.g. needles and syringes, lancets, tourniquets, tubes, sharps container, and swabs), testing devices (i.e. rapid tests), equipment for specimen storage (e.g., cooler box) and transport.
- other supplies: tables, chairs, tents, lights, pencils and pens, paper, folders, waste bin.

Step 3: Identify the recruitment area and interview and testing area(s), and setup area(s) for conducting interviews and HIV testing

A good recruitment area should have the following characteristics:

- clearly defined. The recruitment area should be a well-defined space where potential participants who enter or cross the space are approached for recruitment. Another option is a "moving line". For indoor venues, such as brothels, bars and some shooting galleries, the recruitment area may be inside or outside the venue.
- specific to the venue. The area should be defined to ensure that only people attending the venue are recruited to participate in the survey.
- easy to manage. The recruitment area should be selected to facilitate the survey team's ability to manage recruitment. In larger venues, avoid areas where venue attendees may cross several times, such as near the toilet or bar. These locations can make it difficult to ensure individuals are counted only once.
- of an appropriate size. The recruitment area should be large enough to ensure that the target sample size for each sampling event is reached, but not so large that the survey team becomes overwhelmed.
  - using a small recruitment area for large venues with many people can be helpful.
  - at small venues or those with few people, it helps to use a large recruitment area to recruit a larger proportion of venue attendees. In such cases the entire venue can act as the recruitment area.
  - when there are many people in the recruitment area (e.g. a busy street corner or the entrance to a bar), it is useful to restrict recruitment to only those individuals who enter or cross the recruitment area from a single direction.
  - during an event, the size of the recruitment area can be adjusted to match changing numbers of attendees. Be sure that doing so does not exclude some people from the opportunity to be sampled.
- accessible to the rest of the survey implementers. The location of the recruitment area should allow the field supervisor (or other survey staff) to effectively direct recruitment. It should also be located close to where interviewing and testing are conducted.
- private. The space for conducting interviews and HIV testing and counselling should be out of view and hearing of others at the venue. The space around this area should be controlled to prevent others from invading the privacy of participants, and to limit distractions. Dividers, curtains or tents can be used to create a temporary space for these activities. Venues without suitable space for these activities should not be included in the sampling frame.

Step 4: Determine the measure of size for the event (counting)

As described in Section A-9.6, sampling the target population during a sampling event is similar to second-stage cluster sampling. By this point, investigators have identified the sampling events (similar to first-stage cluster sampling) and now are selecting individuals from each cluster or sampling event. In cluster-based sampling, investigators generally know the number of eligible people in each cluster, but in TLS they do not know beforehand. Therefore, investigators must determine the measure of size (MoS) for an event, that is, the number of eligible people at a venue during a sampling event. The MoS is used to calculate the probability of a person being approached and recruited into the survey. During survey planning, investigators determined the number of venue–day–times from which they can sample. Now, they count the number of people to sample from at a specific venue–day–time.

During each sampling event, a member of the survey staff should be the counter and count venue attendees with a counting device. The same staff member should be the counter for an entire recruitment event to ensure accurate tracking of venue attendees and to avoid counting individuals more than once.

There are many ways to approach counting. One of the most common methods for determining the MoS is to count the number of people at a venue before the sampling event begins, and then count the number that enter the venue during the event. Whichever method is chosen, survey staff should implement it accurately and consistently at all venues in the sampling frame.

Step 5: Hold pre-event staff meeting

Survey teams should arrive at a venue before it becomes crowded in order to set up for the sampling event. This strategy facilitates the pre-event count, planning and set up for the event. Before the sampling event begins, hold a meeting to:

- discuss roles and responsibilities;
- distribute supplies to team members;
- review survey ID, venue code and event number information;

- identify the recruitment area;
- identify the counting area;
- identify the interviewing and HIV testing areas;
- discuss observations about the venue and venue attendees;
- build enthusiasm and raise the energy level of the staff;
- discuss potential safety concerns; and
- conduct other activities as appropriate.

B-1.3.3 Counting

Where to count

The location of counting, or counting area, and the method of counting depend on whether a venue has a dedicated entrance (e.g. a doorway, gate or similar structural entry); counting is conducted differently if the venue does not have a dedicated entrance (as discussed below).

People the counter should not count include:

- Individuals who do not meet the approach criteria. The approach criteria describe who, among all the venue attendees, staff should approach for participation in the survey. These criteria are most useful in venues where not all those present are part of the target population. For example, if a survey's eligibility criteria include MSM aged 15 years and over, males who look much younger than 15 years should not be approached.
- Venue employees or others workers who are not "attending" the venue, and thus do not meet the approach criteria. For example, the male bartender at an MSM bar should not be recruited into an MSM survey, even if he is an MSM. This criterion does not apply for brothels or other venues where the employees are the target population. Other people who may have to enter the recruiting area for work include police officers and street cleaners.
- Individuals who have already been counted.
- Self-referrals, who are people who deliberately enter or cross the counting area trying to enrol in the survey or obtain the incentive. They may have learned about the survey from another venue attendee, or they may be attracted by the activity generated by the survey team. Some individuals will ask if they can participate. They should not be allowed to participate nor should they be counted.

Counting in a venue with an entrance

Two counts are obtained for recruitment events conducted at venues with an entrance: the pre-event count and the entry count.

For the pre-event count, the counter counts all members of the target population who are already inside the venue immediately before the sampling event begins. For the entry count, the counter counts all members of the target population who enter the venue for the first time during the recruitment event. Together, these two counts represent the number of target population members who attended the venue during the sampling event (the MoS). The two counts should be recorded separately and totaled at the end of the event.

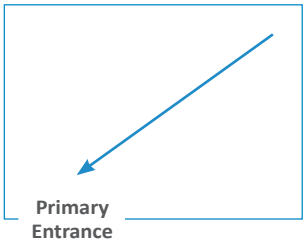
Pre-event count

The pre-event count inside the venue should be conducted even if recruitment, interviewing or HIV testing occurs outside the venue. The counter must personally conduct all pre-event counts; pre-event counts should not be obtained from venue gatekeepers, managers or staff.

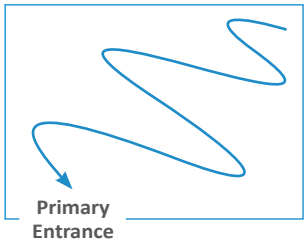
Survey teams should arrive at a venue before it becomes crowded in order to set up for the sampling event before many people arrive. This approach makes the pre-event count and planning and set up for the event easier.

To obtain the pre-event count, the counter should start counting target population members at the point farthest away from the primary entrance that most attendees use to enter the venue. Starting from this farthest point, the counter should count as they move towards the primary entrance, ending the count at the primary entrance. For example, the counter could begin in the back of the bar and then count while walking towards the doorway at the front of the bar (see Figure B-1.1). By starting at the farthest point from the primary entrance, the counter will be able to capture any individuals who enter the venue during the pre-event count. This technique is useful even in smaller one-room venues. The following examples illustrate how to conduct the count, starting with smaller venues (see Figure B-1.1), venues with obstacles (see Figure B-1.2), and progressing to venues with multiple rooms (see Figure B-1.3).

Figure B-1.1 Venues with one room

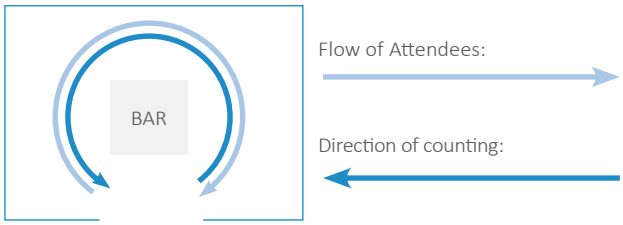


a) The counter should start at the point farthest away from the primary entrance and count as they move toward the primary entrance.



b) To count all the individuals in a crowded venue, it may be necessary to move crisscross through the crowd

Figure B-1.2 Venues with obstacles



If there is an obstacle in a room, such as a large bar, that prevents the counter from moving directly across the room toward the primary entrance, the counter should count while moving around the obstacle and against the flow of venue attendees. By moving against the flow of venue attendees, the counter will minimize the number of individuals who are missed and not counted.[2]

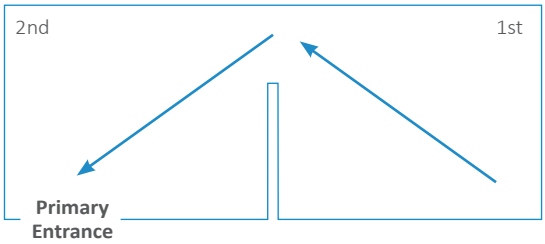
Source: CDC 2014 (*2*)

If there are multiple rooms in a venue, the counter should begin counting in the room farthest away from the room with the primary entrance (see Figure B-1.3). If there are multiple levels and the primary entrance is on the lowest level, the counter should begin counting on the level farthest away from the level with the primary entrance and then count through to the entrance.

Counting in a venue with a complex layout

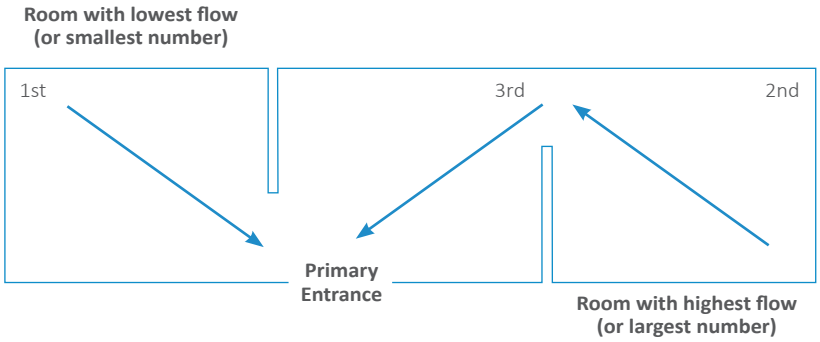
When the rooms or levels in a venue are not arranged in a sequence that ends with the room with the primary entrance, the counter should decide where to start counting based on the flow of people in and out of each room, or on and off each level. The counter should begin counting in the room or on the level with the lowest flow of people and move to the area with the highest flow, and then end in the room with the primary entrance (see Figures B-1.4 and B-1.5). If there is no clear pattern to the flow in the venue, the counter should decide where to start counting based on the number of the target population in each room or on each level. People in the room with the primary entrance should always be counted last.

Figure B-1.3 Venues with multiple rooms



If there are multiple rooms in a venue, the counter should begin counting in the room farthest away from the room with the primary entrance and end counting in the room with the primary entrance.

Figure B-1.4 Venues with multiple rooms not in sequence



[2] Source; CDC 2014: DHAP NHBS Operations Manual.

Figure B-1.5 Venues with multiple floors not in sequence

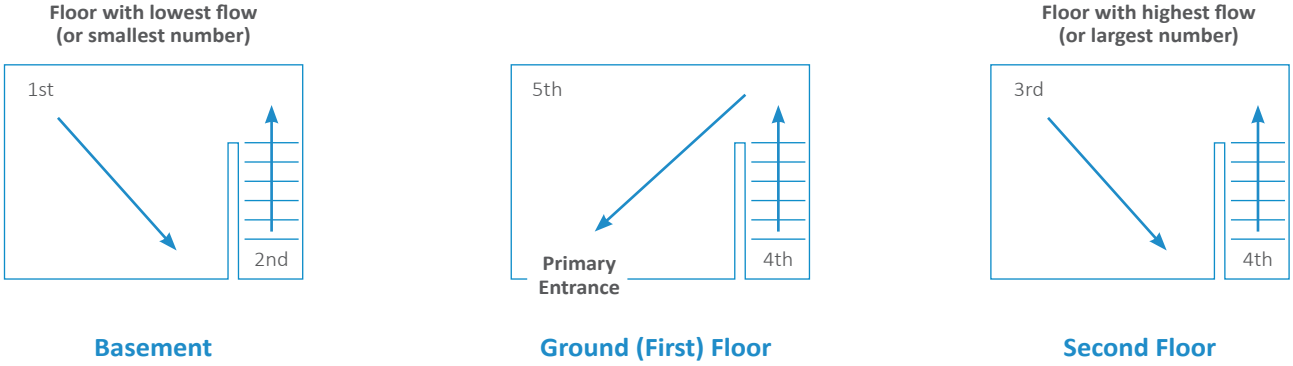
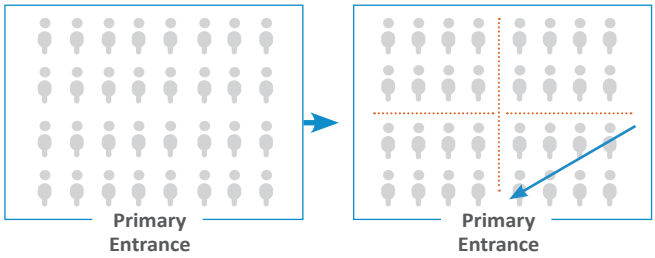
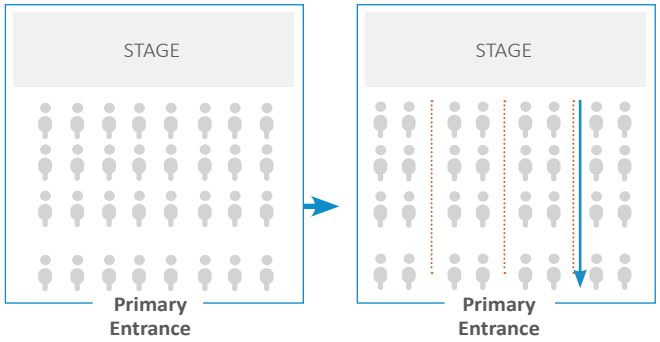


Figure B-1.6 Crowded venues that have equally distributed attendees



If venue is so crowded that it would be difficult to obtain an accurate count, the counter may divide the venue into equal sections, count all the individuals in one of the sections, and then multiply this count by the number of sections to estimate the total count. When the individuals attending the venue are equally distributed throughout the venue, the counter just needs to ensure that the sections selected are the same size. In this example, the counter divides the venue into quadrants, counts the individuals in one quadrant, and then multiplies this count by four to estimate the total count for the venue.

Figure B-1.7 Crowded venues that do *not* have equally distributed attendees



If the counter decides to divide a crowded venue into sections, but the individuals attending the venue are not equally distributed throughout the venue, the counter must take this into consideration. In addition to ensuring that the sections the venue is divided into are the same size, the counter must also ensure that the distribution of individuals is the same in each section. The counter therefore divides the venue into four sections that are perpendicular to the stage to ensure that the distribution of individuals is the same in each section. The counter counts the individuals in one of the sections and then multiplies this count by four to estimate the total count for the venue. Note that the counter moves toward the primary entrance along the line that divides the section. This allows the counter to keep track of the boundary of the section and makes it easier for them to count because they have to count only the individuals between themselves and the wall.

Counting in a venue with a large number of people

In some cases, the number of people attending a venue may be so large at the start of a sampling event that it is difficult to obtain the pre-event count. If this occurs, the counter may divide the venue into equal sections, count all the members of the target population in one of the sections, and then multiply this count by the number of sections to estimate the total pre-event count (see Figure B-1.6 and Figure B-1.7). For example, the counter could divide a busy bar into quadrants, count the people in one quadrant, and then multiply this count by four to estimate the pre-event count for the entire bar. This method should only be used when the entire venue is extremely crowded and it would otherwise be impossible to obtain the pre-event count.

Venues with low attendance or few target population members

Some venues may have very low attendance and hence few members of the target population enter the recruitment area. In this case, an alternative approach plan would be to consecutively approach people seated at a bar, standing along a wall, or already in the recruitment area. If survey staff establish an alternative approach plan for a venue, all staff must follow the plan and not arbitrarily recruit people.

Entry count

The entry count is the number of potentially eligible members of the target population who enter the venue during the sampling event. This count should be obtained at the primary entrance to the venue. The same counter who obtained the pre-event count should obtain the entry count. When the survey staff are ready to begin recruitment, the counter should clear the tally counter to zero and start counting target population members who enter the venue. The counter should stop counting when the last person is approached for recruitment or the field supervisor decides to end the event, usually at the end of the venue–day–time period. Counting should be uninterrupted between these starting and ending points. The counter should continue to count even when all the interviewers are busy with participants.

Some venues may have multiple entrances. If the counter cannot monitor multiple entrances simultaneously, the counter should only count people entering at the primary entrance. The field supervisor should document if a venue has additional entrances where people were not counted.

Counting in a venue without an entrance

Venues without a dedicated entrance such as a doorway or gate include street blocks or parks. If these venues are small enough to conduct a pre-event count, the counting should follow the same steps as those for venues with an entrance.

Many venues without an entrance are so big that it may be difficult to count everyone inside before the sampling event begins. In this case, only collect the entry count. The counter should count individuals entering the recruitment area defined by the survey team.

The methods for counting at venues without a dedicated entrance are similar to those with an entrance. When the survey staff are ready to begin recruitment, the counter should start counting members of the target population who enter the recruitment area. The counter should not count those who are already in the recruitment area when counting begins. If a person is in the recruitment area when counting begins, leaves the recruitment area, and re-enters later during the recruitment event, the person should be counted upon re-entry. The field supervisor should also note that the pre-event count was not obtained because the venue did not have a dedicated entrance.

If there are many places for individuals to enter the venue, count people who enter only from a specific place. The counter should stop counting when the last person is approached for recruitment or, if enrolment is slow, the field supervisor decides to end the event. Counting should be uninterrupted between these starting and ending points. The field supervisor should document when a venue has additional entrances where people were not counted.

Another method is to divide the venue into equal segments of manageable size and conduct the pre-event and entry counts for only one segment. At the end of the event, the survey team should multiply the counts by the number of segments to determine the MoS for the entire venue. For example, if a venue is split into four segments, the MoS of one segment should be multiplied by four to produce the count for the entire venue. This is similar to what is shown in Figure B-1.6. This approach is only applicable if all segments would have approximately equal counts. If one segment of a park has many more people than another, multiplying the count of a given segment would not produce a valid count for the entire venue.

Approach criteria and eligibility screening

Not all venue attendees are part of the target population. The approach criteria describe who the staff should approach to determine whether the person is eligible to participate. Approach criteria should be broad enough to not exclude eligible individuals. However, venues may have multiple populations present. Immediately asking a person if they sell sex may cause problems for the survey team if the person does not sell sex and is unaware of the survey. Thus recruiters should not directly reveal the survey's target population. For example, in a survey of female sex workers (FSW), when approaching a potential participant, the recruiter should say that she is doing a survey of women who go to the venue. Then the recruiter can tell the person how she will be compensated for her time and ask if she would be interested in joining. To determine eligibility while also being discreet in case the person is not an FSW, the eligibility screener should include questions that ascertain eligibility as well as a few questions that hide the purpose of the survey. Such questions might include:

- are you married?
- do you have children?
- where do you meet men?
- how do you earn money?
- have you had sex in the past 6 months?

B-1.3.4 Recruitment and survey steps

Start the entry count after the set-up and the pre-event count have been completed and all field staff are in place and ready to start. Then begin recruiting, interviewing, counselling, collecting and testing specimens as outlined below. Steps 7–9 below may occur in a different order.

Provided recruiters and interviewers are available, recruiters should approach venue attendees consecutively. That is, once the first attendee has been counted and approached, the second attendee should be counted and approached. Recruitment continues while interviewers are available. If there are no interviewers available, counting continues but approaching and recruiting temporarily stop until an interviewer becomes available, at which point they resume. The counter must continue counting during this time to enable investigators to calculate the selection probability of an individual.

Step 1: Approach and recruit
When recruitment begins, the counter or field supervisor directs a recruiter to approach and recruit a specific venue attendee who meets the approach criteria. This process is called the "approach" and can be done by either a recruiter or an interviewer who carries an intercept form (Figure B-1.8). Each recruiter should carry one form. To ensure that recruiters do not preferentially select friends or a certain type of person for participation, the counter or supervisor can identify who should be approached. Recruiters may select individuals themselves provided all have an equal chance of being approached. The counter can direct recruiters at venues without an entrance. At venues with an entrance, the counter can direct recruiters only when the recruitment area is near to the counter, that is, the primary entrance or near the primary entrance the primary entrance to the venue.

The approach begins with the recruiter engaging the targeted venue attendee in conversation. If the person ignores the recruiter, the recruiter should mark a tick next to "No" on Line 1 of the intercept form (see Figure B-1.8 and text box). If the person is willing to talk, the recruiter should briefly describe the survey to the prospective participant and determine whether they have previously participated in this survey round. If the person has already participated, the recruiter should offer thanks for the person's time and mark a tick next to "Yes" on Line 2 of the form before returning to the counter to receive a new person to approach. If the person has not participated in the current survey round, the recruiter should ask to screen the person for eligibility. If the person is not willing to be screened for eligibility, mark a tick next to "No" on Line 3 of the form. If the person is willing to be screened for eligibility, mark a tick next to "Yes" on Line 4 of the form. Each person intercepted should be represented by only one tick. The number of ticks on Line 4 should equal the number of eligibility screening records conducted by the recruiter during the sample event. The number of ticks on Line 5 should equal the sum of all the ticks on the form, which in turn equals the number of people the recruiter approached during the sampling event. See Appendix I-19 for an example of the TLS intercept form and how to summarize the information from these forms after the sampling event is completed.

Figure B-1.8 shows an example of an approach form. Each recruiter should use one approach form for a given sampling event. The form should be used to record recruitment efforts for all individuals approached.

Figure B-1.8 Intercept form

Event Data		
Date:		
Event #:		
Venue ID:		
Recruiter ID:		
Line	**Intercept data**	**Sum**
	Accepted Intercept?	
A	No:	
	Previous Participant?	
B	Yes:	
	Recruited to be Screened for Eligibility?	
C	No:	
D	Yes:	
E	Total:	

If five refusals to approach or to participate occur in succession, the team should stop counting, analyse the situation together, determine the cause of the problem, and develop a plan to resolve the issue. The field supervisor should also evaluate the recruiter's performance and provide recommendations. Potential solutions include seeking advice and assistance from the gatekeeper and trying different recruitment approaches (e.g. more aggressive, less aggressive, using a recruiter of a different gender or ethnicity). After a solution has been identified, counting and recruitment can resume. If the event remains unsuccessful, the team may consider terminating the event. However, all possible solutions to the problem should be tried first.

The number of interviews conducted during each sampling event will depend primarily on attendance at the venue and the number of field staff. The number of interviews conducted during a sampling event at a high-attendance venue will usually be much higher than the number of interviews conducted at a low-attendance venue. To prevent the total sample from being dominated by attendees at a few well-attended venues, investigators should set a maximum number of persons that can be interviewed at one event. Counting and recruitment should continue until the end of the sampling event unless this maximum is reached.

If the count remains low by the end of the sampling event, determine the reason for the low count (e.g. TG do not go to the bar that night or there is a temporary curfew due to civil unrest). If multiple venues have low counts, consider alternative venues and sampling with replacement.

The results from the intercept form are used to calculate the level of response. The MoS is used to weight the data to produce population estimates of results. These topics are discussed in more detail in Chapter C-1.

Sample completed approach form

In this example, 17 people were approached at the sampling event (equals the total number of ticks) by Recruiter 4. Three people refused to speak to the recruiter and one person had already participated in the survey. Two people refused to be screened for eligibility and 11 agreed to be screened.

	Approach form	
	Event data	
	Date Jan 14, 2017 Event # 3 Venue ID B001 Recruiter ID 4	
Line	Participant data	
	Accepted approach	
1	No	|||
	Previous participant	
2	Yes	|
	Recruited to be screened for eligibility	
3	No	||
4	Yes	~~||||~~ ~~||||~~ |
5	TOTAL PEOPLE	17

Step 2: Assign a survey ID
All those who agree to be screened for eligibility (those on Line 4 of the approach form) should be assigned a preprinted survey ID.

Step 3: Screen for eligibility
Screen participants in a private area (see Section A-8 for more information about eligibility criteria). Either the recruiter or the interviewer can screen for eligibility. If the prospective participant is not eligible for participation, the recruiter should offer thanks for the person's time, escort the person from the survey area, and return to the counter to be assigned another person to approach.

Step 4: Obtain informed consent
After a person has been identified as eligible, the interviewer should obtain informed consent according to the protocol and SOP.

Step 5: Interview the participant
The interviewer should interview the participant or supervise a self-administered interview.

Step 6: Counselling and testing
Following the interview, the interviewer should take the participant to the laboratory area to receive pre-test counselling and provide biological specimens. Perform biological testing and relevant counselling according to the protocol, local testing policies and SOP. Ensure that the result of every test (assay) in the algorithm is recorded.

Step 7: Provide reimbursement and prevention material
Field staff can use the time while rapid tests are running to compensate participants and provide them with material on HIV and STI prevention (e.g. informational pamphlets, condoms, lubricants, clean injection kits).

Step 8: Provide test results, post-test counselling, and treatment
If rapid tests are used to test for STI, participants who test positive for treatable STI should be provided with treatment, where possible. If some tests are performed or confirmed off-site, follow protocols to ensure the participant receives test results and treatment later.

Step 9: Provide referrals as appropriate
Refer newly diagnosed and out-of-care participants infected with HIV for care and treatment. All individuals testing negative for HIV should be given information about where they can be tested in the future. Individuals in need of other health or social services should also receive referrals to the appropriate service providers in the community.

Figure B-1.9 illustrates these key steps in survey implementation.

B-1.3.5 Supervision in venues

Strong supervision is crucial during each sampling event. Knowing what occurs during intercepts, monitoring trends in recruitment (refusals, successes), and working with each field-staff member's strengths and weaknesses are critical components of good field supervision. Investigators should observe field staff in a systematic way and provide them feedback on their performance during both training and survey implementation.

Other supervisory activities during the sampling event include:

- Ensure the survey is implemented according to the protocol and SOP.
- Check in with recruiters and suggest ways to improve recruitment techniques (especially responses to recruitment barriers), increase participation and troubleshoot difficult intercepts.
- Monitor recruitment and enrolment throughout the event to determine individual and team performance. Make changes to the counting area, recruitment techniques, team operations or other procedures if necessary.
- Maximize the strengths of survey team members. Determine the best recruiters by observing the quality of the recruitment. Determine who works best at which venues, and with which populations.
- Build team morale. Recognize a job well done and encourage the team to help and support one another.

B-1.3.6 Sampling event closeout

The sampling event finishes either at the end of the venue–day–time period or when the maximum number of participants has been enrolled.

Post-event debriefing

At the end of each sampling event, a post-event debriefing allows field staff to discuss any issues or special circumstances that occurred. This debriefing can be either a meeting convened with all field staff or one-on-one conversations. Some potential questions for discussion:

- in general, how did the sampling event go?
- were there any site-related issues that affected project activities?
- were there any barriers to recruitment? What strategies were successful?
- were there any unusual events (e.g. a participant ended the survey early, or a participant initially consented to HIV testing but had a change of mind)?
- were there any problems with the electronic devices?
- were there any errors with the survey data?
- were there any problems with HIV test specimen collection or test kits?
- were there difficulties with HIV testing and counselling?

If necessary, the field supervisor can record this information in the forms in Appendix I-20, and notify the principal investigator or other staff if the SOP or protocol need to be amended.

Sampling event notes

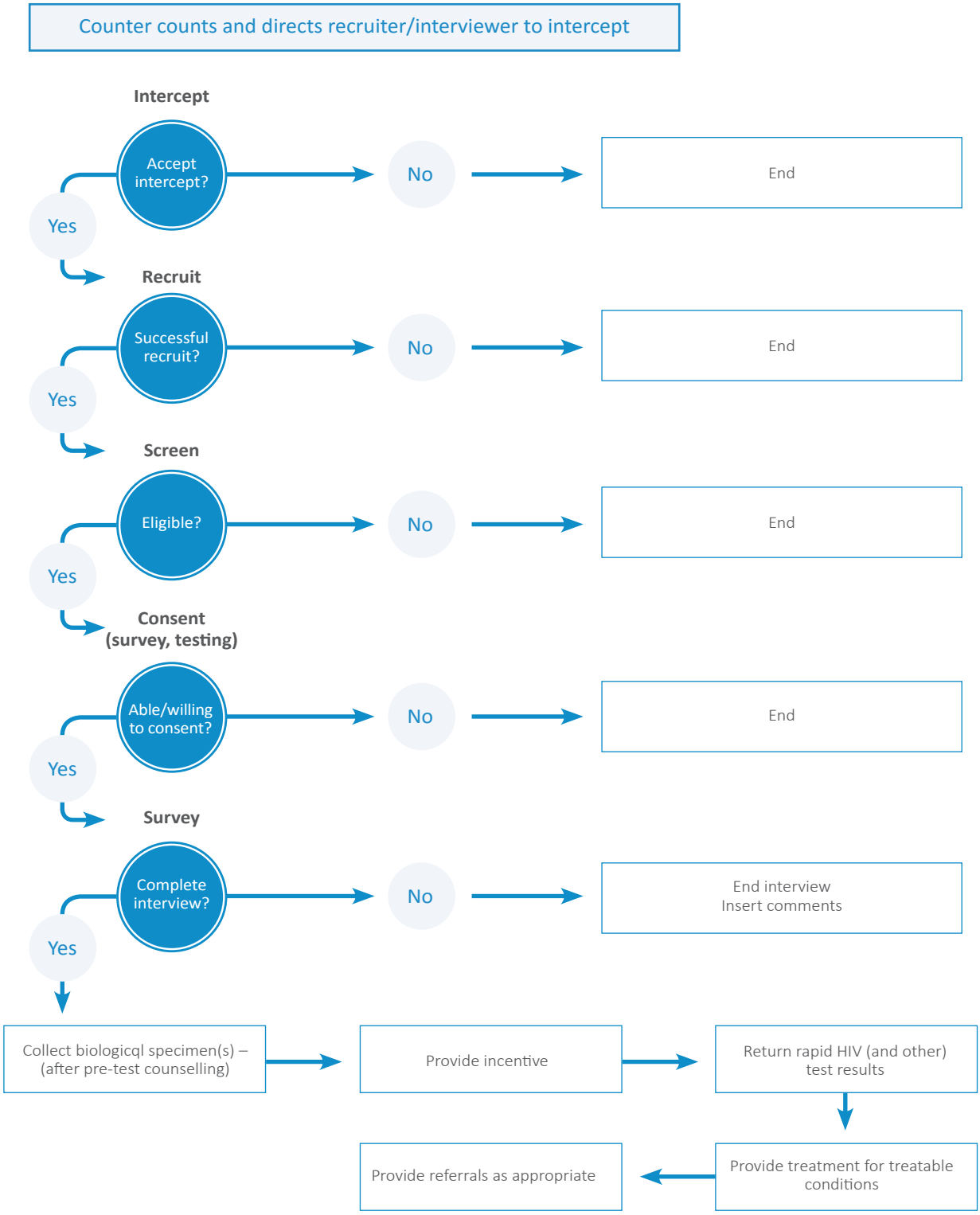
The field supervisor should record notes about the sampling event. These notes document any barriers to survey activities and changes at the venue during the sampling event. This information is important when interpreting recruitment statistics and provides a framework for improving operations. Field supervisors should consider collecting the following information:

- description of the counting, intercepts, recruitment, interviewing and testing areas at the venue;
- barriers to survey activities or safety issues at the venue;
- significant changes in the demographic characteristics or eligibility of venue attendees since the initial venue observation was conducted;
- new venues or day–time periods that were suggested during the sampling event;
- reasons for removing the venue from the sampling frame; and
- information about parking or transportation.

Review of forms and logs

At the end of the sampling event, the field supervisor should collect all materials used in the sampling event from field staff, review them for accuracy, make corrections when necessary, and tabulate recruitment outcomes. Be sure to thank any gatekeepers or others who made sampling possible at the venue.

Figure B-1.9 Biobehavioural survey venue recruitment and survey participation steps



B-1.4 Respondent-driven sampling

Participant recruitment, enrolment, and participation in respondent-driven sampling (RDS) surveys is a complex and multistep process. This section outlines these and other key steps in the implementation and monitoring of RDS surveys.

Participation in an RDS survey usually requires two visits to a survey site. The first visit includes coupon verification, eligibility screening, informed consent, interview administration, biological specimen collection and testing, recruitment training, and compensation for time and transport. The second visit includes return of any remaining test results, recruitment effort interview, and compensation for successful peer referral.

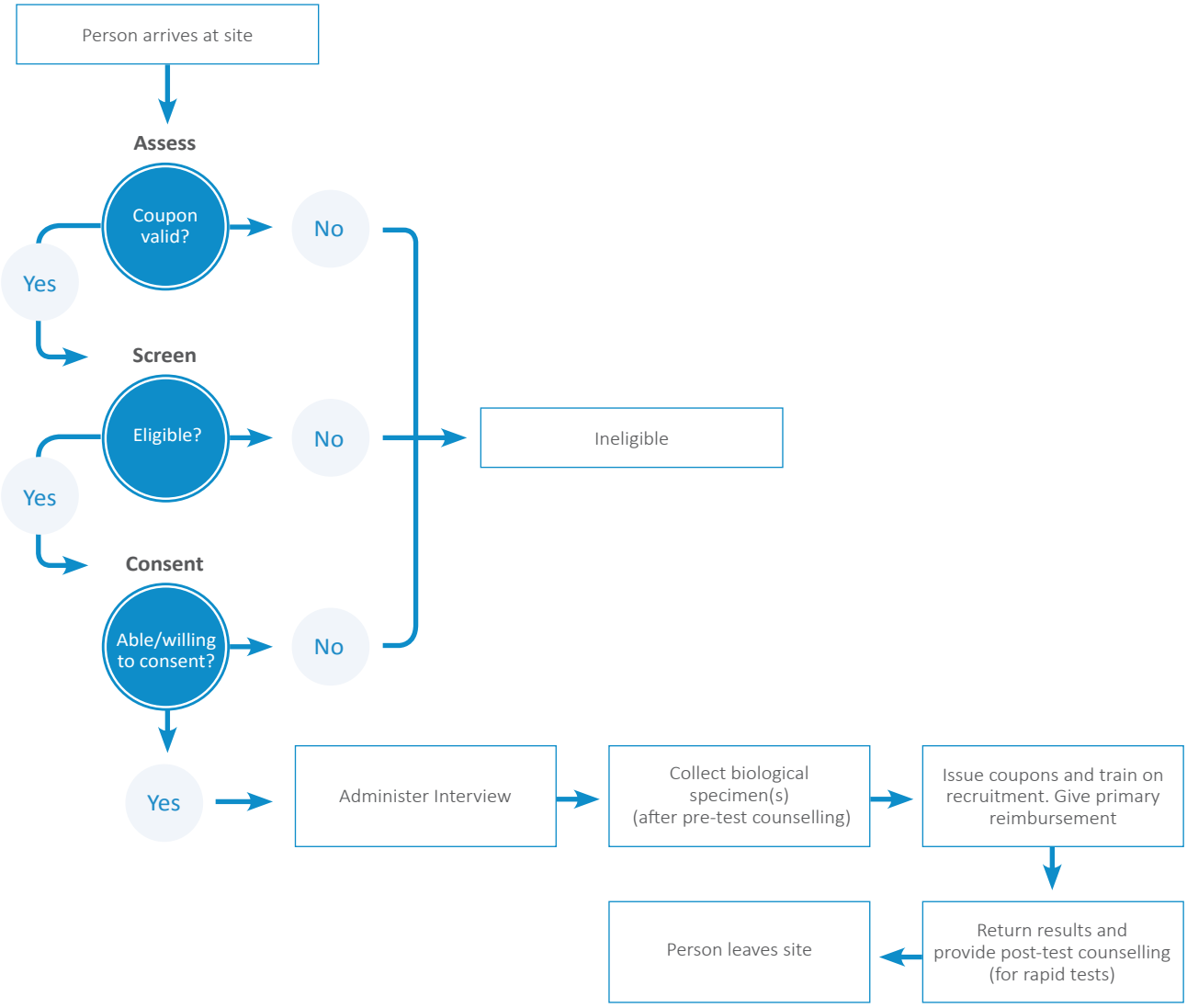
B-1.4.1 RDS survey staff and process

The survey team should include a combination of the following staff positions: field staff coordinator, receptionist, eligibility screener, coupon manager, interviewer, counselor and laboratory technician. For details regarding the staff involved in an RDS survey, see Chapter A-14.

First visit in an RDS survey
Because the first and second visits for an RDS survey involve multiple steps, checklists are recommended to ensure that all steps have been performed for each participant (see Appendix I-32).

Figure B-1.9 shows survey procedures for the first visit.

Figure B-1.9 Procedures for the first visit for an RDS survey



Coupon verification

Potential participants are greeted by a receptionist upon entering the survey site. The receptionist verifies that each participant:

- has a valid coupon (except for seeds, for whom a coupon is not required); and
- is able to provide informed consent (i.e. is not drunk or under the influence of drugs, and has no mental or emotional disabilities that may preclude informed consent).

A valid coupon is a coupon that has not been forged or photocopied. If a start date, expiry date, or both are included, the potential participant must generally participate during the specified period (see Section A-9.7), although the investigators may choose to relax the rules to encourage participation. Doing so would not violate any RDS assumptions. The participant should be given an information sheet describing the survey while waiting for the next step in the survey process: eligibility screening. A sole receptionist is often the first person who can identify repeat participants: because the receptionist will see everyone who enters the survey site. Individuals who have already participated in the survey must be refused participation. Their coupons should be marked "VOID" and kept by the survey team.

Survey staff should treat all participants in a courteous manner; participants who feel dissatisfied or offended may refuse to recruit peers. Also, participants who have a bad experience may tell others about their experience, which can further reduce participation.

Ensure that the waiting area is a pleasant place with, for example, a television or reading materials (including health-related materials) to keep visitors occupied. Population-specific materials should be avoided to protect participants in the event of unexpected visitors. Investigators should decide whether family members (e.g. children of SW) or friends are permitted in the waiting area. Although some participants may feel uncomfortable if people who are not members of the target group are in the waiting area, allowing others to be present might also help reduce barriers to participation. Formative assessment can help identify the best approach to this decision.

Eligibility screening

The first step in eligibility screening is to verify the coupon's validity in the survey database, which is often done by the coupon manager. If the coupon is valid, the manager should stamp it "USED". If the coupon is invalid (i.e. copied, edited or already used), the manager should stamp the coupon "VOID" and file it for record-keeping. All coupons should be kept for record-keeping and to prevent reuse. Individuals bringing invalid coupons are not eligible to participate in the survey and must be refused participation. Their coupons should similarly be collected, stamped "VOID" and filed.

Eligibility screening often involves asking sensitive questions and should occur in a private room or area. A key component of the screening is to determine whether the potential participant is a member of the target population. If possible, the screener (i.e. the person responsible for the eligibility screening, often the coupon manager) should have experience working with the target population, or be a member of the target population. Screeners can also verify membership through further probing. For example, PWID usually have visible injection marks. If a potential participant has no visible injection marks, the screener can ask how drugs are prepared and injected, or the cost of the drugs. Only individuals who actually inject drugs will be able to answer these questions correctly. Experts on the local context can help tailor questions to identify the target population. Identification of members is harder with SW and MSM, because they can be part of the target population but not know the main areas where sex is sold or where men meet each other. Membership verification questions for each key population are included in the questionnaire module for eligibility in Section II of the supplementary materials.

Sometimes, potential participants try to enrol by pretending to be a member of the target population, especially if compensation for participation is high. Ineligible individuals should be asked to leave the survey site, especially if they are ineligible because they are not part of the target population. Staff must still be polite to these individuals. If using an electronic data-collection tool, consider saying "I'm sorry, but the computer indicates you are not eligible", which reduces the chance of protest.

Informed consent

After verifying eligibility and before conducting any other survey procedures, survey staff must obtain informed consent. A member of the survey staff reviews the consent form with the participant and answers any questions. Depending on the protocol, the participant can either sign the informed consent form (written informed consent) or verbally agree (verbal informed consent). The staff member should sign the document after informed consent is obtained. Many institutional review boards (IRBs) allow the use of verbal informed consent for surveys of key populations in order to facilitate participant anonymity. See Chapter A-3 for more information on informed consent. After informed consent is obtained, the participant is enrolled in the survey and the interview to determine the size of their personal network can start. If informed consent is not provided,

if appropriate, ask the person the reason for declining to participate in the survey. Document the reason and provide the person with transportation compensation before escorting them from the survey site.

Size of the personal network

Questions on the size of the personal network are essential for calculating weights for RDS population estimates. These questions should be administered face-to-face by the coupon manager or interviewer to ensure participants understand the questions correctly. Size of the personal network is determined by a series of questions that lead to an estimated total number of people the participant knows who are probably eligible for participation and have been seen during a defined period. See Section A-9.7 for examples of these questions.

Interview

The questionnaire should be administered to all enrolled participants, including seeds. After the interview, the participant proceeds to another area for counselling and to provide biological specimens for testing.

Specimen collection

Participants should receive counselling before they provide biological specimens. Participants must provide consent for the interview and collection of biological specimens. Consent for these two distinct activities may be obtained together or separately. Participants may consent to the interview but not to testing and specimen collection. Reasons for refusing to provide biological specimens may include initial fear of or unfamiliarity with the survey and the process. Some participants may provide informed consent for specimen collection but later change their mind. Any change in the informed consent should be noted and respected. Asking such participants why they changed their minds is acceptable in order to respond to their concerns and improve procedures for future participants. For example, participants may decline an HIV test if they already know they are infected; the staff member can explain to that testing will help the survey to determine the needs of everyone in the population. The survey may also offer additional tests for those living with HIV, including CD4 and viral load.

Coupon issuance, peer-recruitment training and primary compensation

After the interview and collection of biological specimens, the participant is given coupons to recruit peers into the survey. Each coupon has a unique code. The participant's own coupon is linked to the coupons they give to peers. The link between the coupons can be tracked by RDS Coupon Manager or by using a logbook (see Section A-9.7 for more information on RDS coupon design and distribution).

At this step in the process, the role of coupons is explained to participants. Survey staff train participants on how to recruit peers using the following instructions:

- safeguard the coupons.
- give coupons only to people you know who are part of the target population.
- give only one coupon per person.
- inform peers of the purpose of the survey and what the survey entails, for example, completing a questionnaire and providing blood for tests for HIV and STI testing.
- return to the survey site for the second visit to be compensated for successful recruitment, that is, for all of your recruits who completed the interview (as a minimum).

To ensure consistency in training recruiters, provide a written script for staff members administering the training (see Appendix I-31). A script minimizes selection bias from inconsistent recruitment training.

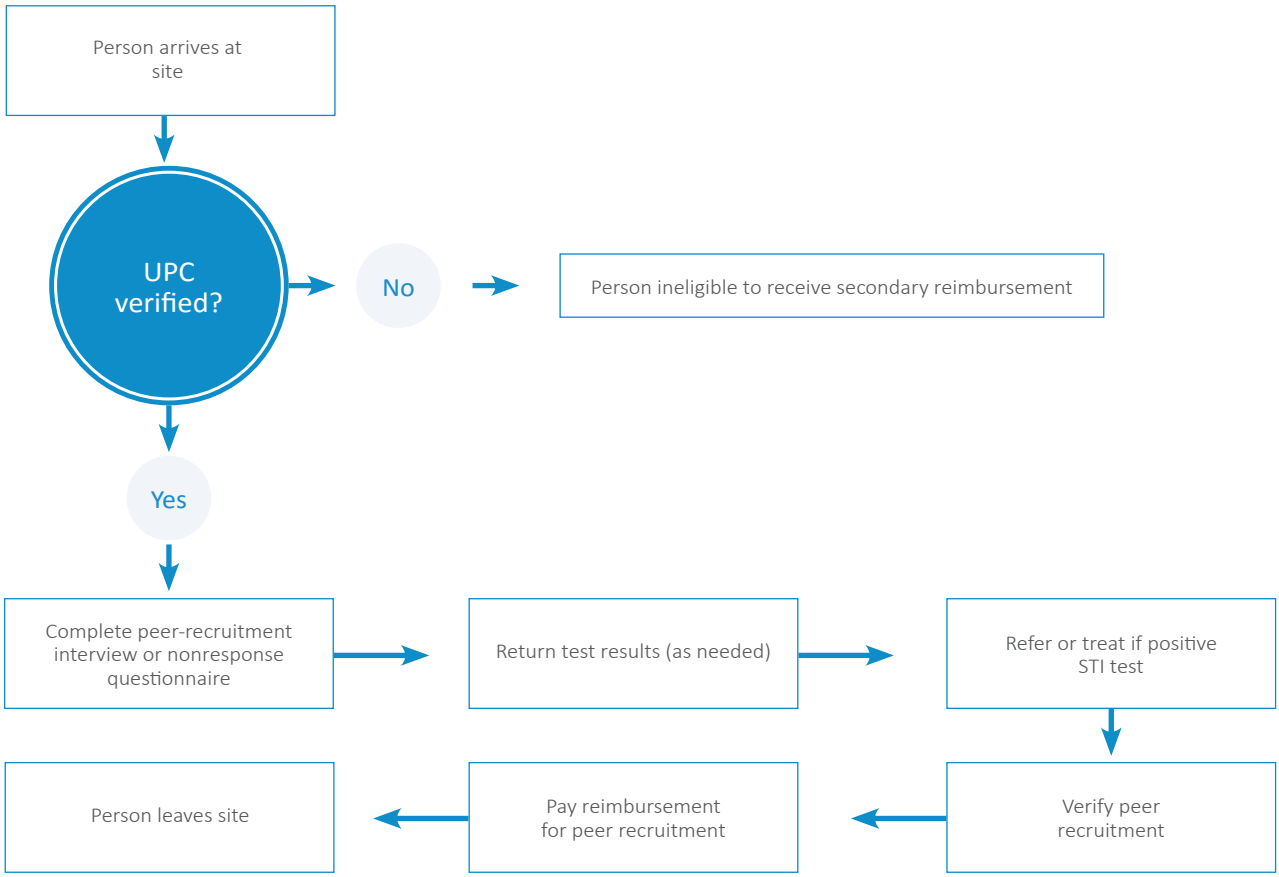
The unique participant code (UPC) may also be created at this point (see Chapter A-9) to verify participant identity at the second visit. Participants should be encouraged to collect their recruitment compensation during the second visit, when they may also receive results for any non-point-of-care tests. A second visit should be scheduled, usually 2 weeks after the first visit. If coupons include an expiry date, the second visit should be scheduled after this date to encourage recruits to return before the recruiter's second visit.

The last activity of this step is to provide compensation for participation. Primary compensation should be provided only to those who complete at least the interview, and should be recorded in a database or logbook. Also, provide the participant with risk-reduction materials such as condoms, lubricants, syringes and needles, as well as educational materials on HIV and other STIs.

Issues with compensation

Participants may contact survey staff to learn which of their recruits have participated in the survey, in order to encourage others to participate and thus gain more compensation. Others may ask at their second visit. Giving the coupon numbers of those who participated to the recruiter may jeopardize participant confidentiality and is not recommended. Further, recruiters may place undue pressure on those who do not participate, and responding to such requests can add logistical challenges for survey staff because the person contacted usually does not have recruitment information.

Figure B-1.10 Second visit in an RDS survey



Return of test results

If rapid testing is part of the survey, results should be returned to the participant towards the end of the first visit. Post-test counselling and treatment, if feasible and appropriate, should also be provided. Link participants who test positive for HIV and those testing positive for other STIs to appropriate care and treatment. If tests require processing at an off-site laboratory, participants will need to return for their test results. Results should be returned during the participant's second visit.

Second visit in an RDS survey

The purpose of the second visit is to compensate a participant for recruiting peers, to conduct a short interview about peer-recruitment efforts, and to return test results that were processed off-site. Figure B-1.10 shows the steps in a participant's second visit.

Verification of participants' identity

The second visit begins with verification of participants' identity, for which participants are asked to show their appointment reminder card with their survey ID. If a UPC has been created (see Section A-9.7), the series of questions is asked again to recreate and confirm the personalized code. This procedure can also be used to find participants' survey ID if they lose it. Fingerprint scanning can substitute for a UPC.

Secondary compensation

Verify whether any of the participant's coupons have been returned by their recruits. Compensation should be provided only for recruits who joined the survey and completed the interview.

Peer recruitment or nonresponse interview

In most other sampling methods, investigators can easily estimate the nonresponse rate because survey staff directly sample and enrol participants. In RDS, however, measuring nonresponse is difficult because participants

do the sampling on behalf of the investigators. There are three reasons for unsuccessful recruitment:

- the participant lost the coupons or did not offer coupons to peers.
- the participant offered coupons to peers but the coupons were rejected.
- the coupons were accepted by peers who did not attend the survey site.

The peer recruitment questionnaire (see Section II-40 for an example) can be used to determine patterns in coupon distribution, acceptance and refusal. Because not all participants will return for the second visit, results from this questionnaire should be interpreted with caution.

Return of test results

If testing was conducted off-site, test results should be returned during the second visit, with STI counselling and treatment provided if appropriate and feasible. If treatment is not available on-site, participants should be referred to target population–friendly clinics for treatment.

B-1.4.2 Recruitment monitoring

Because the participants conduct sampling in RDS, predicting when the target sample size will be reached is difficult. In addition to sample size, the analytical and functional assumptions of RDS (see Section A-9.7) must be met in order for the sample to produce valid population estimates. Recruitment monitoring throughout the survey process helps investigators identify and address issues in recruitment.

Sample size

The time taken to reach the target sample size depends mostly on the actual sample size chosen, the number of coupons given to each participant, the number of seeds used, the amount of compensation, the density of the survey population's networks, and the willingness of individuals to join the survey. Although the investigators cannot control the size of the target population, other factors can be controlled by them, including accessibility of the survey site, the days and hours of operation, the number of participants the survey team can enrol per day, seed quality, and the quality of participants' experience. Investigators have a number of tools available to increase recruitment.

Increase the number of coupons

The number of coupons issued per participant can be modified throughout the recruitment period. The maximum number of coupons issued can be increased either temporarily or permanently in order to boost recruitment (e.g. from 3 to 5 coupons). The number of coupons given to each participant and the date of this procedural change should be documented. The increase in coupons issued per participant should ideally be temporary, lasting only long enough to boost recruitment. Successful RDS recruitment includes long recruitment chains. If too many coupons are issued to each participant, shorter chains with more branches may result.

Add more seeds

The number of seeds investigators can add is unlimited. Adding more seeds can boost recruitment, especially if new seeds can reach untapped social networks. However, adding more seeds can also reduce the chances of producing the long recruitment chains required in RDS.

Increase compensation

Increasing compensation can further encourage the target population to participate in the survey and recruit peers. However, a large increase may lead participants to sell their coupons or to lie about their eligibility or their membership in the target population in order to receive the compensation. Excessive compensation may also be coercive and unethical.

Conduct continued formative assessment

Investigators can solicit feedback on how to improve sampling from participants and key informants. For example, through informal conversations with participants and key informants, investigators may learn that the site location is not accessible to the entire target population and therefore an additional site is needed; or that perhaps site hours need to be changed. An anonymous suggestion box can also be placed in the waiting or reception room to encourage feedback.

Other aspects

In addition to monitoring the sample size, other aspects of sampling require monitoring, including the survey response rate, sample composition, number of recruitment waves, convergence, bottlenecks, and cross recruitment between sites.

A coupon-management system is required to track who recruited whom, ensure that compensation is paid to the correct person, and track the number of waves reached. If on-site computers are available, coupons can be tracked in an electronic database. A commonly used database is RDS Coupon Manager[2] (see Section A-9.7).

[2] www.respondentdrivensampling.org

Survey response rate

The nonresponse rate in an RDS survey cannot be accurately measured because investigators depend on participants to refer potential participants. The investigators have little knowledge of those who were offered a coupon but chose not to participate, and how these people differ from those who came to the survey office. Nevertheless, nonresponse can be approximated by monitoring the following:

- number of coupons issued and redeemed. Track how many coupons have been issued to survey participants. A coupon is considered "redeemed" when a person with a valid coupon is screened for survey eligibility.
- peer recruitment interview. The peer recruitment questionnaire (see Section II-40) is administered when participants return for a second visit. This questionnaire can determine how many of each participants' coupons were offered, accepted and refused. Monitor the proportion of second visits to determine whether those who return for a second visit are similar to those who do not return.

Monitoring sample composition

Because the external validity (generalizability) of RDS surveys is difficult to determine, it is good practice to compare sample characteristics with formative assessment findings. Such characteristics include age, sex, ethnicity, neighbourhood of residence and HIV status. If the sample composition during recruitment does not approximate that of the target population, additional seeds with the missing characteristics can be included in an attempt to increase their representation. For example, if a formative assessment revealed that both male and female individuals who use drugs were found in the survey area, the sample composition should comprise both males and females. If only males were recruited, female seeds should be added to encourage recruitment among females.

Number of recruitment waves, equilibrium and convergence

The number of recruitment waves by seed should be monitored weekly. Equilibrium and convergence should also be assessed throughout recruitment. The two indicate whether the sample is independent of the characteristics of the seeds. While equilibrium is based on the sample, convergence is based on the population estimate. Having diverse seeds can help achieve equilibrium and convergence faster. Variation in both sample and population characteristics will be limited after they are reached. Continually perform, for example, bottleneck and convergence plots for key variables (these plots can be created in software such as RDS Analyst[3]) to determine whether equilibrium and convergence have been reached. If they have not been reached for key variables by the time the sample size is reached, continue enrolment if funding and the protocol allow. Seek ethical approval if the protocol needs to be revised for a larger sample size.

Multiple survey sites

If the survey has multiple survey sites, separate networks can form around each survey site. This violates the RDS assumption that the survey population is made up of a single network. Cross recruitment between sites must occur (i.e. participants from one site must recruit people who go to another site) in order to prevent violation. Cross recruitment is easily monitored if coupon numbers identify where the participant was recruited (see Section A-9.7). To encourage cross recruitment, investigators should inform participants that their recruits can go to any survey site.

B-1.4.3 Ending an RDS survey

As the survey approaches the target sample size, investigators need to end sampling in a controlled manner. The number of coupons issued can be gradually reduced or stopped altogether. The goal is to ensure that the sample size is reached while limiting the number of people who may seek participation after this point. To the extent possible, participants presenting with valid coupons should continue to be accepted and enrolled. If additional people can no longer participate (e.g. the IRB-approved sample size has been reached or the budget has been exhausted) survey staff should politely explain the situation to those seeking participation. A script should be developed to guide survey staff in communicating this message to potential participants. During recruitment training, survey staff can also advise participants that their recruits should come to the survey site quickly before the sample size is reached. As survey sites will remain open for scheduled second visits, eligible participants with valid coupons presenting themselves during this time should be reimbursed for their transportation, even if not enrolled.

[3] http://hpmrg.org/software/

2. Survey monitoring

This chapter describes considerations for survey monitoring. Monitoring is the real-time assessment of survey implementation to ensure that the protocol and standard operating procedures (SOPs) are followed, and to identify challenges while the survey is being implemented so that immediate corrective action can be taken. Monitoring is not a one-time event but an ongoing activity throughout the survey. Monitoring key aspects of survey implementation can contribute to improved data quality and should be done by survey staff and investigators.

Table B-2.1 lists key areas for monitoring. It may be useful to create a simple checklist based on Table B-2.1 to systematically document the monitoring of each activity, including what was monitored, by whom and when; any problems or challenges identified; and how each problem was addressed. Documenting monitoring activities can be useful not only for improving the existing survey, but also for future survey development and for when an evaluation is conducted after the survey (see Chapter C-3).

Table B-2.1 Monitoring topics and areas of focus

Topic	Areas of focus
Staff	• Adequate number and type of survey staff (e.g. interviewers, nurses or counsellors, data managers, field staff supervisors)
Training and training material	• Field staff trained on SOPs to ensure protocol adherence
Sampling	• Sampling conducted properly according to protocol and SOP
Recruitment	• Verification of duplicate enrollees • People posing as members of target population • Fake/invalid coupons, coupon uptake (RDS) • Inappropriate reimbursement, selling of coupons (RDS)
Enrolment	• Participant consent administered properly (see Supervision section below) • Any respondents enrolled that were actually ineligible • Ineligible people received coupons (RDS) • Respondent return rates for follow-up activities including return of test results and necessary referrals
Data collection	• Adequate spaces for interview administration • Missing interviews • Revision of questions which: - are difficult to understand - lead to refusals - lead to extreme values • If the questionnaire is modified for any reason, field staff should conduct mock interviews to validate any changes in electronic data entry, including verifying new skip patterns, and valid ranges for new variables

Topic	Areas of focus
Data management	• Adherence to data-management SOPs including data storage, security, backup (see Section C-3)
Data quality	• Data-quality checks stated in protocol and SOPs conducted to identify and address the following: - missing data - invalid data values - outliers - skip pattern errors - refusal rates
Laboratory	• Adequate space for specimen collection • Adequate equipment • Supervision of biological specimen collection and processing for transportation to the testing laboratory • Record-keeping (e.g. specimen quality form or specimen tracking form) consistent with protocol and SOP for both survey laboratory and/or referral laboratory • Trained laboratory staff perform testing • Temperature logs completed and maintained • Test kits, reagents, and specimens not expired and being stored at appropriate temperatures • Verification of the number of specimens at laboratory and the number of completed interviews with the number of participants who provided consent for interview and/or specimen collection • Proportion of test results returned, delayed • Missing specimens, poor quality specimens • Poor testing quality • Insufficient supply of test kits • Testing errors • Inclusion of external quality assessment to validate results (see Chapter A-6)
HIV counselling, referral and reimbursement	• Conduct of HIV counselling and testing • Provision of relevant referrals and reimbursement
Human subjects	• Breaches in confidentiality • Informed consent procedures followed
Participant burden and acceptance	• Time spent for survey participation • Aborted interviews or survey procedures
Data-collection costs	• Actual data-collection costs compared with estimated budgeted costs

B-2.1 Observing data collection

Field supervisors should observe about 10% of the total number of interview sessions. Participants must give verbal permission before the session observation. Field supervisors should ensure that privacy is maintained during the interview. More frequent observations (about 25% of the observed interviews) may occur towards the beginning of the survey to identify any potential problems or issues in the administration of the survey. The field supervisor may use a site supervision checklist to identify potential issues. Problems should be addressed as soon as possible.

For qualitative interviews, field supervisors should verify that interviewers have sufficiently probed key themes. If qualitative interviews are conducted with audio recording, the interviewers should also listen to a sample of audio recorded interviews to ensure that the recording quality is adequate for future transcription.

Feedback from survey participants may be obtained either immediately following survey participation by randomly approaching participants, or using a locked suggestion box where participants may leave anonymous notes about concerns they may have.

B-2.2 Laboratory monitoring

As with interviews, field supervisors should select and observe about 10% of specimen collections. More frequent observations (at least 25% of specimen collections) may occur at the beginning of the survey. Field supervisors should conduct daily or more frequent

checks of specimens to identify and correct potential problems with specimen quality, including:

- haemolyzed blood specimens
- improperly dried blood spots
- illegible specimen labels
- missing labels
- inadequate specimen storage

For quality control and assurance, the laboratory may be enrolled in a proficiency programme relevant to assays performed there. In addition, a proportion of specimens may be sent to a reference laboratory or external laboratory for retesting to assess consistency and quality of results. All assay runs should include testing of valid controls and calibrators, if applicable. Any invalid runs should be repeated. Concordance between initial and confirmatory runs should be assessed as an additional measure of quality.

Participant follow-up and referral

Field supervisors should check that test results are available for post-test counselling and referral. For sites that do not provide same-day test results, supervisors should monitor logs to assess the proportion of participants who return for their results. If the proportion of participants who do not return for their results is high (e.g. greater than 20%), survey investigators should actively encourage participants to return for test results during pre-test counselling sessions. Please refer to Chapter A-6 on biomarker considerations for further information.

Section B

3. Preparing data for analysis

This chapter provides guidance on data cleaning, preparing data for analysis, and assessing data quality. It discusses the different types of errors commonly found in surveys, and provides information on how to assess the quality of survey data, address identified errors, append and merge datasets, and document any modifications to the data.

There are several steps in preparing a dataset for analysis. Data cleaning (Steps 1–3) should ideally occur throughout data collection. Merging and preparing datasets for analysis (Steps 4–5) should occur after all data have been collected and cleaned.

B-3.1 Identifying and correcting data errors

Data cleaning is the process of detecting and addressing data errors in a dataset. Data cleaning can be minimized if survey investigators take steps to assure data quality throughout data collection (see Chapter A-13). If electronic data collection is used, errors due to data entry can be minimized by proper coding, correct skip patterns, and correct data-validation rules. In addition, the survey instrument should be thoroughly tested before implementation and data collection should be monitored closely, to find and correct repeated mistakes early. Data managers should maintain a log showing corrections that are made. Some software programs automatically track corrections. If corrections are made on paper before data entry, then notes and the initials of the person making the changes should be placed on the paper form.

Data errors can originate from various sources, including data collection, data entry (e.g. transcription errors, creating duplicate data) or sampling. Using a systematic process to identify and address errors can improve the quality of survey data.

The following are examples of common data errors and concerns:

- transcription errors
- duplicate records
- measurement error
- lack of internal consistency
- out-of-range values
- outliers
- other errors

These errors are described in Table B-3.1.

Table B-3.1 Common data errors and how to prevent them

Common data errors
Transcription errors
Description of error: Transcription errors are mistakes made when a person enters data from one form of documentation to another, for example, when paper-based data are entered into an electronic database.
How to prevent error: Double-data entry can minimize errors of transcription. Two survey staff members each enter the data once and examine values that do not match within the two versions of the dataset. Many programs used for data entry will include functions and internal checks that can identify such discrepancies.
How to resolve error: Each error identified in the dataset should be verified with the original response from the paper-based questionnaire. After data verification, all transcription errors should be corrected and any changes or corrections documented for future reference.
Duplicate records
Description of error: Duplicate records occur when one participant has more than one record in a dataset. The two records are the same for all variables for this participant and can result from transcription errors or errors in data merging.
How to identify error: Duplicate records can be identified by running frequencies of the survey ID variables to identify duplication, during either the data-entry process or data management, including uploading of surveys and merging with other data sources.
How to resolve error: For truly duplicate records (all variables for two or more records are identical, including the survey ID), delete the extra records.
Measurement error
Description of error: Measurement error occurs when the response provided differs from the true value. It can be caused by: *Questionnaire design:* The wording of questions can affect the responses provided by the participants, for example, if questions are unclear, ambiguous or suggestive. *Interview administration:* The mode of administration – face-to-face, computer-assisted personal interview (CAPI), audio computer-assisted self-interview (ACASI). For example, participants may not answer questions of a sensitive nature in front of an interviewer but may be more comfortable responding through a computer. Alternatively, they may not fully understand a question asked in ACASI and may not ask survey staff for clarification. *Interview:* The interviewer can introduce error in survey responses by not reading the question or response options as intended, by probing inappropriately, or by adding other information that deviates from the question or confuses the participant. *Participant:* Participants can interpret the meaning of, or respond to, questions differently. Errors can also be due to recall (i.e. the participant cannot accurately remember the true answer).
How to prevent error: Measurement error can be reduced through piloting of the data-collection instruments before survey implementation and obtaining feedback from participants on their understanding of the questions and response options (i.e. cognitive interviewing). If measurement errors are found during survey implementation, the data collection instrument can be modified (ensuring that the version change of the instrument is documented). Additionally, if face-to-face or CAPI is used for data collection, interviewers should be trained thoroughly in order to not influence participants. Uniformity in interview administration is essential.
How to resolve error: Measurement errors can be difficult to detect, particularly if the responses are not out of range or are consistent with responses to other variables in the dataset. If measurement errors are detected, keeping the original values and documenting the source of the error may be the best option.
Lack of internal consistency
Description of error: Internal consistency means answers to different questions do not contradict each other. Answers that do contradict each other are called logic errors and indicate a lack of internal consistency. Internal inconsistency can result from poor questionnaire design, incorrectly defined skip patterns, incorrectly programmed skip patterns, or a combination of these factors. For example, a participant states being male and later in the interview states being pregnant. A skip should have been programmed to prevent males from responding to any questions directed towards females, such as being pregnant. Or participants may incorrectly respond to a question; for example, reporting always using a condom in the past 6 months but indicating in a different question that they did not use a condom at last sex. Sometimes logic checks can prevent such errors; for example, a logic check can be placed in an electronic questionnaire to require that age at marriage is no more than current age.

Common data errors
How to prevent error: Most errors in logic are preventable when validation rules and skip patterns are programmed correctly (see Chapter A-13). Investigators should pilot data-collection instruments before implementing the survey to ensure the tools are programmed correctly. If errors are found during data collection, the tool should be revised or reprogrammed as appropriate, with changes documented.
How to resolve error: In some cases, inconsistent responses can be coded as missing. However, some errors due to incomplete data or faulty skip patterns cannot be resolved. In this event, document the error and be aware of its implications for data analysis or presentation of results.
Out-of-range values
Description of error: Out-of-range values are values that do not make logical sense. An example of an out-of-range value is a participant response or record of "15" for month of birth when the range for the variable should represent the number of the month (i.e. 1–12). An invalid value would include values outside the given response choices. For example, if a question requires a response of "1" for YES or "2" for NO and the response value of "3" is given, the error must be corrected and set to "Missing" because it is an invalid response.
How to prevent error: Programming minimum and maximum values for ranges and valid values into an electronic data collection tool can minimize out-of-range and invalid values.
How to resolve error: Errors due to out-of-range and invalid values can be resolved in three ways. First, set the value to "Missing". Second, for continuous variables, use either the mean or the median value for the entire population, depending on the distribution of the data; for data that are not normally distributed, using the median is preferable because it minimizes the impact of outliers. Third, infer the correct value through other responses or data sources, if possible. Each change and the reason for the change should be documented.
Outliers
Description of error: Outliers, or extreme values, are not always errors. Outliers can be defined statistically. Formal statistical tests can identify outliers. These tests are designed to identify values that may influence statistical analyses practically. For an outlier to be considered an outlier, the value is illogical. For example, to the question "How many sexual partners have you had in the past week?" a response of "100 partners" may be possible, but is illogical. Typing errors may be suspected of producing some outlier values; for example, a response of "33" partners in the past 6 months could have resulted from typing the number 3 twice, when the participant actually responded "3".
How to prevent error: Because outliers or extreme values are not always errors, there are very few ways to prevent them. Questions can be worded carefully to elicit more consistent responses from participants. Interviewers can be instructed, or ACASI/CASI can be programmed, to ask the question again in a nonjudgemental manner to ensure the participant understood the question. Additionally, outliers or extreme values that do not make practical sense can be minimized by programming ranges into the data-collection tools.
How to resolve error: Some experts recommend that these values remain "as is" in the analysis because they are not necessarily errors (*3*). However, the investigator may want to consider the influence of these extreme values on the analysis before deciding whether to include the data in the final analysis. The analysis can be repeated with and without the outlier data to assess the impact of the outlier on the analysis. After observing the outlier's effect, and if these values are set to "Missing", they should be reported as excluded from the analysis. The effect of outliers and extreme values can sometimes be minimized through categorization of response values. For example, a question probing the number of noncommercial sex partners in the past 6 months may yield a few outliers such as "250", when all other values are below, for instance, 50. Placing all the continuous values in categories denoting 0, 1, 2 or 3+ noncommercial sex partners would minimize the effect of outliers.
Other errors
Other errors that require decision rules for correction include missing forms or records, programming errors, and values or records that have been incorrectly changed or deleted during data handling or cleaning. Although these errors cannot be corrected, they must be documented and noted in any reporting of the analysis.

Step 1: Assessing data quality

The first step is to assess the quality of the survey data by examining the data for the common errors mentioned above. Screening data can help identify missing data, errors such as out-of-range and invalid values, outliers and unusual patterns due to measurement errors or lack of internal consistency. Methods for data-quality assessment can be observational and not necessarily statistical. Invalid values, ranges and logic checks should be determined before beginning the data-quality analysis. Some commonly used methods for assessing data quality are described below (see Table B-3.2).

Table B-3.2 Methods of screening data for errors

Method	Description	Pros and cons	Level of difficulty
Frequency distribution	Visually check for errors to identify missing data, out-of-range values, invalid values and other potential errors.	**Pros:** • high level of skill not required for observation **Cons:** • may not identify all errors due to subjective review • rules defined by data-cleaning plan may not be consistently applied across all versions of a dataset.	Easy
Bivariate display or 2 × 2 table	Bivariate displays or cross-tabulations look at expected relationships among variables. Particularly helpful to find logic errors.	**Pros:** • high level of skill not required for observation • can verify internal consistency or logic checks **Cons:** • some analysis skill required to create 2 × 2 tables • may not identify all errors due to subjective review • rules defined by data-cleaning plan may not be consistently applied across all versions of a dataset.	Medium

An automated programme can be used to flag errors. The programme should be created using the rules defined in the data-cleaning plan, including logic checks, valid ranges and values, and missing values. Such programmes objectively screen all data and can be consistently and repeatedly applied to all versions of a dataset. However, their creation requires knowledge of syntax for statistical software and, often, significant time, due to the level of detail needed if the programme is based on the defined rules.

Examples of using the data-quality assessment methods described in Table B-3.2 are provided below.

Example 1: Transcription error and missing data
Observing the frequency distribution of a variable can help to identify missing or duplicated data. For example, by monitoring the frequencies of the survey IDs, the number of participants can be confirmed. If the survey has collected information from 15 participants, but there are only 14 participants according to the frequency distribution (Figure B-3.1), investigators should determine why a record is missing and which record is missing. The data could be truly missing, or, as shown in Figure B-3.1, one individual may have mistakenly been given the survey ID of another participant.

A frequency distribution of survey IDs can also help determine duplicates. A survey ID should be unique; therefore, the frequency of every survey ID should not be greater than 1 (see survey ID 8 in the table below). If the data is duplicated, go back through the original data and forms if possible to determine the correct value for the variable.

Consider whether a variable has too many missing values. Generally, if more than 10% of the values of a variable are missing (and the missing values are not due to skip patterns), the variable may have too many missing values to be interpreted.

Figure B-3.1 Example of data table with transcription error

Survey ID	Frequency	Percentage	Cumulative percentage
1	1	7.14	7.14
2	1	7.14	14.28
3	1	7.14	21.42
4	1	7.14	28.56
5	1	7.14	35.70
6	1	7.14	42.84
7	1	7.14	49.98
8	**2**	14.29	64.26
9	1	7.14	71.40
10	1	7.14	78.54
11	1	7.14	85.68
12	1	7.14	92.82
13	1	7.14	100.00
Total	**14**	100.00	

Example 2: Out-of-range values and invalid values
Out-of-range values and invalid values can be prevented if valid ranges and values are assigned and programmed correctly during the creation of data-collection tools. During data collection, out-of-range and invalid values can be identified using frequency distributions.

For example, for month of birth, the range of valid values should be 1 through 12 to represent each month in the year. In Figure B-3.2, the value 13 is out-of-range and should be corrected based on the rules in the data-cleaning plan. The electronic programme for the data-collection tool should also be modified to prevent future out-of-range values for this variable.

Figure B-3.2 Example of data table with out-of-range value

Month of birth	Frequency	Percentage	Cumulative percentage
1	5	12.50	12.50
2	2	5.00	17.50
3	2	5.00	22.50
4	1	2.50	25.00
5	3	7.50	32.50
6	2	5.00	37.50
7	3	7.50	45.00
8	4	10.00	55.00
9	7	17.50	72.50
10	3	7.50	80.00
11	4	10.00	90.00
12	3	7.50	97.50
13	**1**	**2.50**	**100.00**
Total	40	100.00	

Figure B-3.3 Example of data table with other potential error

HIV test results	Frequency	Percentage	Cumulative percentage
Negative	122	81.33	81.33
Positive	**26**	17.33	98.66
Indeterminate	2	1.33	100.00
Total	150	100.00	

Example 3: Other potential errors
Frequency distributions can be used to identify other potential errors. For example, if the laboratory identified 27 HIV-positive specimens but the frequency distribution shows only 26 HIV-positive observations (Figure B-3.3), then the correct number requires verification through further investigation.

Example 4: Bivariate displays
Bivariate displays or cross-tabulations can also indicate errors in data collection such as missing values, out-of-range or invalid values, and other potential errors. Bivariate displays can verify internal consistency and perform logic checks in the data. For example, a cross-tabulation of the variables "ever had sex" and "had sex in the past 6 months" should show zero responses in the cell for those who responded "no" to ever having sex and "yes" to having sex in the past 6 months (not "3" as in Figure B-3.4). Changes to the data should be based on the rules in the data-cleaning plan.

Figure B-3.4 Example of bivariate display

Had sex in the past 6 months	Ever had sex		
	No	Yes	Total
No	21	54	75
Yes	**3**	177	180
Total	24	231	255

Step 2: Address identified errors

If errors are detected, the next step is to address the identified errors and missing values based on the rules defined in the data-cleaning plan. As part of this process, the person responsible for resolving the possible error and the period in which the resolution will occur should be specified. Having these rules available to survey staff involved in data collection and management is recommended. Correcting errors may involve returning to the original source of information (i.e. questionnaire or initial dataset). Three solutions exist if the error is in the original data:

- correct the value:
  - use value in the original source of information (for paper-based collection);
  - if continuous, set the value to the mean or median;
  - if possible, locate information from another response or source;
  - let the value stand; or
- set the value to "Missing".

Any changes made to the dataset must be documented in a data-audit log that describes the screened errors and the resolution of the errors.

Recording the reasons for missing values

Missing data can have various effects on the data-set. They can decrease the effective sample size, making estimates less precise (i.e. wider confidence intervals [CIs]) and statistical tests less powerful in detecting any significant differences. Also, missing data may not be randomly missing – that is, certain characteristics may be associated with missing data, which can affect the results of analysis. Some groups of participants may have different proportions of missing data. Therefore, as far as possible, investigators should prevent or minimize missing data. If there are no patterns in the missing data among the different groups, estimates will not be biased by the missing data. If missing data are ignored, those who did not respond to a certain question are assumed to be, on average, similar to those who did respond to the question.

For missing values, imputation can also be considered. Imputation is a statistical method in which values are prescribed for the missing data based on other data provided by the participant. There are various ways to impute data, including multiple imputation, but these methods are complex and require consultation with a statistician.

Variables can have a value of "No response" for any of several reasons, which may be useful to distinguish. A value of "No response" could be due to the following:

- the participant did not know or remember a response to a question;
- the participant refused to answer a question;
- the question was skipped based on a previous question; or
- the participant did not input a response and left a blank space for the question.

Some situations require distinguishing between the missing response types. For certain questions, a response category may mimic the default category "Don't know". For example, a question such as "Can a healthy-looking person have HIV?" has the response options "Yes", "No", "Don't know" and "Refuse to answer". In this example, the response option "Don't know" is a legitimate response and should not be set to "Missing". In the questions below in Figure B-3.5, the first example demonstrates when "Don't know" should be considered a response option. The second example demonstrates when "Don't know" should be set to "Missing."

Ideally, make the numeric code for "Refuse to answer" or "Don't know" the same for each question. In the example below, "8" is the code for "Refuse to answer" each time. If using the same value consistently, ensure that this value cannot be a potentially valid response in one or more questions. For example, do not use "8" for "Refuse to answer" if a question that needs a "Refuse to answer" response option asks for the number of years the individual has lived in the location, because "8" could be a valid response to this question. "888" may be a better choice for "Missing" in this case.

Figure B-3.5 Example of the different meaning of "Don't Know"

Example 1. "Don't know" is a legitimate response.

Can a healthy person have HIV?	YES 1 NO 2 **DON'T KNOW 3** REFUSE TO ANSWER 8

Example 2. "Don't know" is not a legitimate response.

Are you currently living with a sexual partner?	YES 1 NO 2 **DON'T KNOW 7** REFUSE TO ANSWER 8

If applicable, any identified errors that are due to the questionnaire design or programming of the electronic data-collection tool should be corrected. These changes should be documented and a new version of the data-collection tool should be saved.

Step 3: Append and merge datasets

After the data from each dataset have been cleaned and the errors have been corrected, the next step is to append the data (if there are multiple sources) and merge them, as discussed below.

Appending datasets

If the same data are collected through multiple sources (e.g. multiple sites), creating a database by appending the datasets from those sources is highly recommended. Appending a dataset means to directly add the observations from one dataset to another, usually because each dataset is derived from different participants. When investigators append, they are usually adding two datasets measuring the same variables in different groups together, essentially stacking one dataset on top of another. The variables and variable names should be consistent throughout the data sources. In Figure B-3.6, datasets 1 and 2 are appended, resulting in one dataset with four records.

Figure B-3.6 Example of appending datasets

Dataset 1

	Variable 1	Variable 2
Survey ID 1		
Survey ID 2		

Dataset 2

	Variable 1	Variable 2
Survey ID 3		
Survey ID 4		

Appended dataset

	Variable	Variable 2
Survey ID 1		
Survey ID 2		
Survey ID 3		
Survey ID 4		

Figure B-3.7 Example of merging datasets

Dataset 1

	Variable 1	Variable 2
Survey ID 1		
Survey ID 2		

Dataset 2

	Variable 3	Variable 4
Survey ID 1		
Survey ID 2		

Merged dataset

	Variable 1	Variable 2	Variable 3	Variable 4
Survey ID 1				
Survey ID 2				

Merging datasets

If different data are collected from the same participant (e.g. behavioural questionnaire and laboratory results), the data should be merged (Figure B-3.7). Merging must be performed using the key variable in common (usually survey ID), and checked to detect unmatched records from each data source. After the datasets are merged, a selection of records from the merged file should be compared to the original data sources to ensure the data were correctly merged. The data manager should verify that records from each dataset are not missing from the merged set. If there are missing records, the source of the error must be identified and corrected. If a correction is not possible, the team must document the decision on how the missing records were handled. Data managers should ideally merge datasets multiple times throughout the data collection, instead of waiting until all data are collected, in order to ensure that any problems are discovered early and thus can be rectified.

Step 4: Documentation

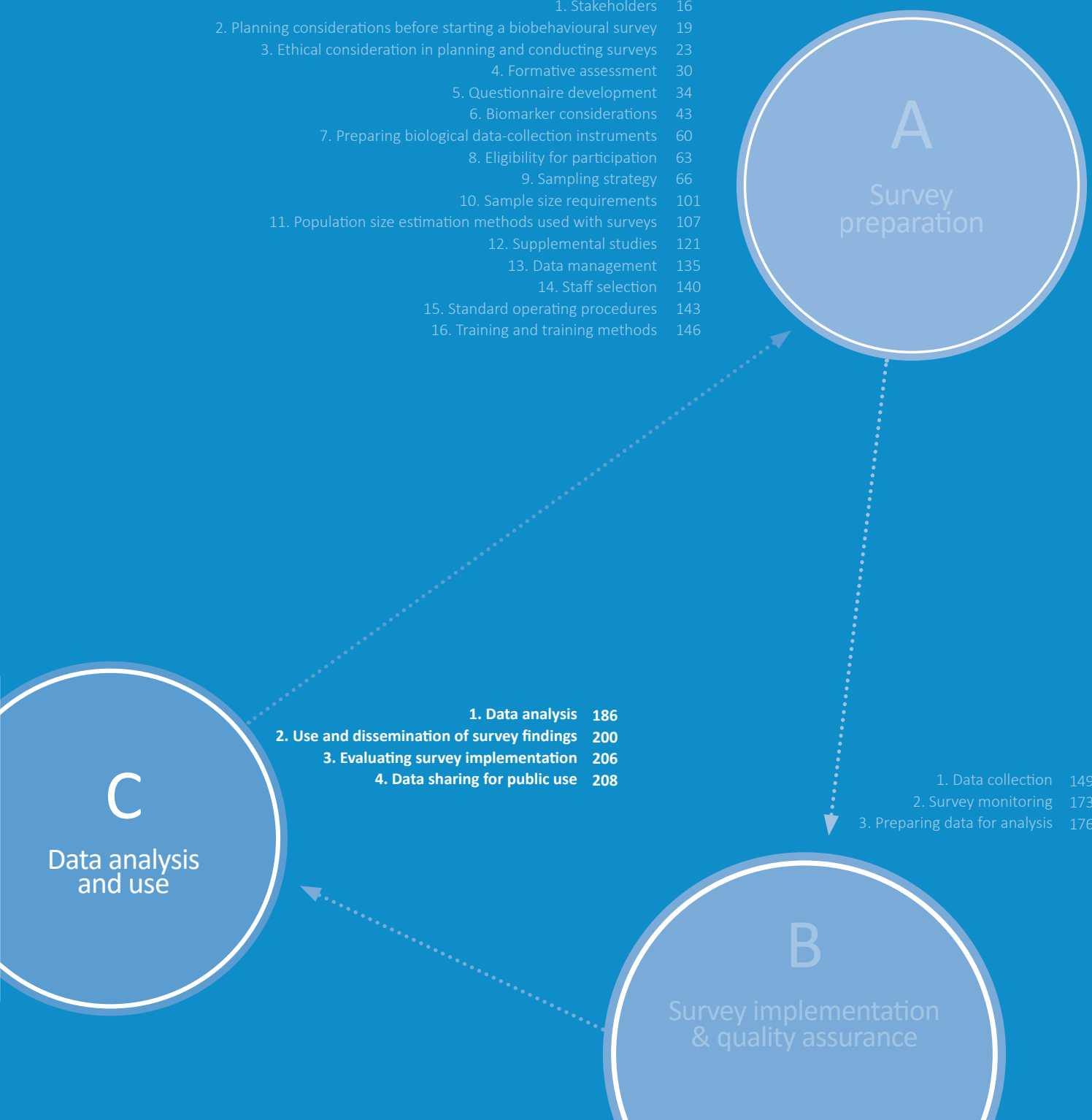
A good data-management practice is proper documentation of all procedures, including any changes made to the data as part of the data-quality assessment. Documentation is important for several reasons. It enables each survey staff member who is working on a dataset to systematically follow all the steps and actions that were taken, such as records deleted, variables recoded, new variables created, and categories combined or collapsed. Further, if the dataset were shared internally or with external partners, others could readily duplicate and verify the analysis if all data preparation steps were properly documented. Documentation is also important if the survey results are to be disseminated and published, especially in a peer-reviewed journal.

The following are suggestions for good documentation practice:

- keep the original dataset separate and save all programs leading to the creation of the final dataset, to allow recreation of any intermediate dataset if needed;
- document in a separate log any modifications made due to data entry;
- document computer programmes used to modify the original dataset through assigning a unique identifier that includes a descriptive title and the date the programme was created. if possible, describe the changes and modifications within the programme itself;
- adopt a system of naming conventions for datasets when they are merged or modified;
- maintain a log of programmes run on the dataset, including when the programme was run, the dataset used, and a brief description of what was run;
- when creating variables through recodification, select new and meaningful names to preserve the original variables;
- if surveys are conducted at multiple sites, variables should be uniformly named and coded across sites and surveys;
- maintain the source files as originals before applying corrections; and
- ensure all variables and categorical values have descriptive labels and are formatted appropriately.

C. Data analysis and use

1. Stakeholders 16
2. Planning considerations before starting a biobehavioural survey 19
3. Ethical consideration in planning and conducting surveys 23
4. Formative assessment 30
5. Questionnaire development 34
6. Biomarker considerations 43
7. Preparing biological data-collection instruments 60
8. Eligibility for participation 63
9. Sampling strategy 66
10. Sample size requirements 101
11. Population size estimation methods used with surveys 107
12. Supplemental studies 121
13. Data management 135
14. Staff selection 140
15. Standard operating procedures 143
16. Training and training methods 146



1. Data collection 149
2. Survey monitoring 173
3. Preparing data for analysis 176

1. Data analysis 186
2. Use and dissemination of survey findings 200
3. Evaluating survey implementation 206
4. Data sharing for public use 208

Section C

1. Data analysis

This chapter provides an overview of how to analyse data from biobehavioural surveys (BBS). This includes using the data analysis plan (developed with the protocol) as a guide in preparing a dataset for analysis, creating sampling weights for BBS data, and reviewing analytical methods for BBS data. The chapter includes only basic analytical methods, such as descriptive statistics that summarize BBS data for survey reports. More advanced statistical methods such as regression analysis are beyond the scope of these guidelines.

The BBS data-analysis strategies presented in this chapter focus on probability-based sampling designs and inferences (see Chapter A-9). Data analysis that fails to account for the complex survey design of BBS may produce inaccurate estimates, confidence intervals (CIs) and *P*-values. Accounting for sampling design during analysis ensures that the CIs (the measure of uncertainty around the estimates) include both random sampling error (the measure of precision of the estimates) and other sources of error, such as refusals or selection bias. The standard errors of complex sampling designs (e.g. cluster sampling) are different from those of simple random sampling (SRS). Briefly, the survey design should aim to maximize the precision of the estimates and minimize the standard error; however, there are many factors that influence decisions on survey design, including budgets. For example, cluster sampling may improve the ease and reduce the costs of surveying the sample, but tends to reduce the precision of the estimates and increase the standard error because participants sampled within clusters are often more similar to one another than to participants from other clusters (cluster effect). The decrease in sampling costs often outweighs the loss of sampling efficiency (i.e. increases in standard error).

Although most statistical software packages can account for complex survey designs, these analyses should be performed by a statistician or other person knowledgeable about the methods the packages use. A summary of software for survey analysis is available online.[1] However, analysis of data from surveys using respondent-driven sampling (RDS) typically requires specialized software such as RDSAT or RDS-A, or specific additional packages to software such as Stata.

C-1-1 Developing a data analysis plan

Data analysis turns raw survey data into actionable information that can be disseminated through presentations, technical reports, peer-reviewed publications, policy briefs, press releases and other dissemination tools (see Chapter C-2). To ensure that the survey objectives are satisfied when the data are analysed, develop a comprehensive data analysis plan along with the survey protocol described in Chapter A-2. A data analysis plan includes the following components:

- survey objectives and research questions to be answered;
- timeline for data analysis;
- roles and responsibilities of those involved with data analysis;
- type of product (e.g. technical reports or manuscripts, or both);
- description of any variable transformations (e.g. grouping continuous variables into categories or creating new variables from a combination of variables);
- sources of data (e.g. other datasets to be used for trend or comparative analysis);
- description of the data-cleaning process;
- strategies for addressing missing data;
- description of the weighting process;
- statistical methods and software to be used;
- data-collection tools; and
- table shells (templates) and a description of how data will be presented and which cross-tabulations will be made (including specific denominator definitions).

[1] http://www.hcp.med.harvard.edu/statistics/survey-soft/

C-1.2 Preparing a dataset for analysis

Chapter A-13 explained processes for transforming raw survey data into a single data file. This section describes the next step; that is, preparing the dataset for specific analyses to address survey objectives outlined in the data analysis plan. Before beginning weighting and analysis, data should be reviewed to assess the need for any transformations and to address the issue of nonresponse.

Creating new variables

Sometimes, it is necessary to transform or recode variables to fulfill survey objectives. Creating new variables in preparation for data analysis may include collapsing multiple response categories, combining multiple variables into one new variable, or categorizing continuous variables. One common situation for data transformation is viral load laboratory results. These have a wide range of possible values, making the results difficult to analyse and interpret. By transforming raw HIV RNA laboratory data into a logarithmic ($\log_{10}$) format, the resulting variable range is narrower, more manageable, and easier to analyse, interpret and graph. A log scale is based on the powers of 10 and is calculated as $\log_{10}$(# copies/mL) or 10^y = # copies/mL. For example, if an HIV RNA laboratory result is 10 000 copies/mL, then $\log_{10}$(10 000), which is equivalent to 10^4 = 10 000. Thus, the transformed variable would be $\log_{10}$ = 4.0, a more manageable value for analysis than 10 000 copies/mL.

Before starting this process of transforming or recoding variables, investigators should consider the number of responses in each resulting stratum (category). A sufficient number of responses in each stratum is important for producing meaningful estimates, especially if the data will be weighted.

Collapsing response categories

New variables can be created by collapsing response categories for existing variables. For example, consider condom use where consistent condom use is defined as "Always uses condoms with a partner". If the three response categories for original condom use with a partner variable are "Always," "Sometimes" and "Never", the responses for "Always" would represent consistent use of condoms, and "Sometimes" and "Never" would be combined to represent not consistently using condoms with a partner. The original variable has three response categories and the new variable has two. Other examples of this process are collapsing the number of age categories (e.g. five age groups down to three age groups) or values for marital status (e.g. never married, currently married, divorced and widowed groups into "Ever married" and "Never married"). Looking at the data may also be helpful to determine how to collapse categories. For example, if two adjacent categories are similar regarding the measure of primary interest, it may be appropriate to collapse these two categories. If the variable is numeric, it may be helpful to collapse according to quantiles, or according to commonly used public-health categories.

Combining two or more variables

New variables can also be created by combining responses from two or more variables into a single variable. For example, a new variable "men who have sex with men (MSM) who sell sex and have never been tested for HIV" would be created by combining the variables for MSM who sell sex and HIV-testing in MSM. Verify that new variables are created correctly by comparing the frequencies of the original and new variables.

Categorizing continuous variables

Sometimes, recoding continuous variables (e.g. age or number of partners) as categorical variables is useful. However, meaningful results depend on carefully defined categories. Categories may be defined according to published literature, biologically or clinically meaningful values, or statistically meaningful points such as a median (i.e. data are divided at the 50th percentile) or quartiles (i.e. data are divided into four equal increments at the 25th, 50th and 75th percentiles). Although categorizing continuous variables may make analysis and interpretation of results easier, when cutoffs are applied to continuous data, there is a loss of information, and a reduction in the statistical power to detect effects *(4)* and the precision of estimates (e.g. means or odds).

To illustrate these limitations, consider a survey of respondents ranging in age from 18 to 54 years. A graph of the original age distribution might reveal two peaks (i.e. a bimodal distribution), one at 24 years and one at 38 years. Important information about age distribution might be lost if the age variable is categorized into two groups defined by the median or into four groups by quartile. The resulting estimates by age group might be misleading because interpretation may suggest a constant effect throughout each age group and an artificial change in effect between categories. In contrast, continuous variables such as CD4 count and viral-load data might be more meaningful if categorized according to clinical guidelines that can be interpreted in terms of disease severity or treatment adherence. Verify that new variables are created correctly by comparing the distributions of the original continuous and new categorical variables.

Addressing nonresponse

Missing or nonresponse data may influence estimates and must be evaluated before analysis begins, particularly with variables known to be correlated with the outcomes of interest. This evaluation should examine the types, patterns and magnitude of nonresponse. Nonresponse is of particular concern when it is not

random (i.e. individuals with certain characteristics are more likely to fail to provide an answer to a question than are other individuals); estimates will be biased if this issue is not addressed in analysis. Several types of nonresponse are possible in surveys:

- Unit nonresponse. Sampled units (e.g. persons or venue) did not participate in the survey. The differences between those who participated and those who chose not to participate may have a significant impact on estimates of HIV prevalence or risk. Post-stratification weighting for unit nonresponse can reduce, but not eliminate, the influence of this bias on the survey results. The process uses a set of external population estimates (e.g. census data) to calculate weights that adjust the sample data to align with key characteristics (e.g. age, sex) of the population. For example, compared with younger MSM, older MSM may be less likely to attend certain venues, or less likely to be recruited into an RDS survey, which may make the population appear younger, on average, than it actually is, or appear to have a lower HIV prevalence than it actually has.
- Item nonresponse. A participant gave no response or an invalid response to at least one survey item. Examining participant characteristics of item nonresponses provides context for interpreting estimates. Nonrandom item nonresponse might indicate problems with poorly constructed questions, response options or skip patterns, or might indicate differential sensitivities to questions. Depending on the size of the survey sample and the objectives of the analysis, item nonresponse might not affect the analysis. For small sample sizes or items that are directly related to analysis objectives, analysts should examine nonresponse patterns by participant characteristics and survey administration (e.g. with electronic data-collection tools). For example, individuals who refuse to be tested for HIV may be more likely to already know that they are HIV positive, which can make the HIV prevalence of the population appear lower than the actual prevalence.

Some statistical software packages have commands that allow for assessment of patterns in missing data (e.g. "mvpatterns" in Stata). This assessment provides analysts with information to help decide which strategies are most appropriate to address the types of missing data in the sample. Weighting (e.g. post-stratification weighting for nonresponse) and imputation – replacing a missing value with an estimated value – are two methods of compensating for missing data during analysis. A sensitivity analysis measures the effect of different methods of dealing with non-response on the final results of the BBS. Several software packages offer options for running sensitivity analyses (e.g. PROC MI in SAS v 9.4). In brief, a sensitivity analysis examines violations of the assumption that data are missing at random by running different models and comparing the consistency of the results of those models. If the results of multiple models produced by the sensitivity analysis are consistent, then it is likely that the strategy to deal with non-response was reasonable and produced reliable BBS results. If the sensitivity analysis produced inconsistent results across the multiple models, it would be necessary to consider a different strategy to address non-response data. These procedures are complex and are beyond the scope of these guidelines; they should be performed by a statistician or other person knowledgeable about the methods.

Analysing and interpreting results from nonprobability surveys

Because the selection probability of participants in nonprobability surveys is unknown, weighting data from such surveys is impossible. Probability-based analyses to assess associations are similarly impossible, because the data are representative of only the survey participants and not of the larger population. Data from such surveys should therefore be interpreted with caution. Findings may be used to describe a subgroup of the population but should not be extrapolated to the entire population.

Creating sampling weights for BBS data

Weighting adjusts the data from survey participants so that the estimates are representative of the population from which the participants were drawn and thus reduces some of the different types of bias that were defined in Chapter A-9. The weights being calculated here are the selection weights only.

As mentioned in the sampling section (Chapter A-9), sometimes oversampling is used to ensure that enough people are selected to obtain stratum-specific estimates large enough to make inferences about that subgroup. In such cases, a design weight should be used to compensate for oversampling (or undersampling). This design weight ensures proportionate representation of the oversampled subgroup in the overall population. The computation of a design weight follows the same process as that used for other sampling weights below.

As described in Chapter A-9, in a probability sample, each individual has a known chance of being selected for participation in a survey. Sampling weights are computed as the inverse of the probability of being sampled:

$$w = 1/P$$

where

w = sampling weight

P = probability of selection

If all participants have been assigned a weight equal to the inverse of their selection probability, the sum of the

weights will be equal to, or very close to, the size of the sample population. For example, if clusters of unequal sizes are selected via SRS, then the weights will probably not be equal to, but should be close to, the size of the sample population.

For most BBS, however, participants have unequal probabilities of being selected, resulting in biased samples. When a sampling design results in unequal selection probabilities for participants, weighting techniques are used to give more weight in the analysis to those who had less chance of being selected. This process improves generalizability or representativeness of the estimates. Nonresponse and post-stratification weights may still be needed.

In general, there will be one final analysis weight for each participant. This final analysis weight is the product of the individual sampling weight and other adjustments as needed (e.g. a nonresponse adjustment factor and a post-stratification adjustment factor, as shown in the example below).

$$w_{final} = w_{selection} \times w_{nonresponse} \times w_{post\text{-}stratification}$$

where

$w_{selection}$ = individual sampling weight ("base weight"), which is the inverse of the probability of being selected. Selection probability may be influenced by more than one factor, and thus might have multiple components. The final $w_{selection}$ is the product of the inverse of all of these components.

$w_{nonresponse}$ = a nonresponse adjustment factor, which is the inverse of the probability (i.e. $P = 1 -$ nonresponse) that a selected individual will participate in the survey.

$w_{post\text{-}stratification}$ = a factor that uses population distribution data (e.g. census data) to adjust the distribution of survey participants.

Weighting for cluster-based sampling: conventional cluster sampling and time-location sampling

Table C-1.1 shows examples of cluster-based sampling stages in sampling designs. See Sections A-9.5 and 9.6 for more information.

An additional stage might be the probability of selecting geographical units, if they are used in the sampling frame; for example, selecting a number of towns from a list of all towns. This would become the first stage and be included in the probability of selection:

$$PI = p_{geographic\ unit\ (if\ applicable)} * p_{time\text{-}location\ cluster} * p_{individual\ venue\ attendee}$$

Step 1: Calculate the sampling probability

There are four common scenarios for cluster-based sampling: conventional cluster sampling (CCS) and time-location sampling (TLS). The following examples provide guidance on calculating sampling probability based on how clusters are selected.

Table C-1.1 Examples of cluster-based sampling stages in sampling designs

	Conventional cluster sampling (CCS)	Time-location sampling (TLS)
First stage (cluster)	Venue	Time-location cluster
Second stage (individual)	Individuals associated with venue	Venue attendees during specified period

Scenario 1:

- First stage: selection of clusters by probability proportional to size (PPS)
- Second stage: equal number of individuals selected from each cluster

$P_i = (m * M_i/M) * (n_i/N_i)$

where

P_i = probability that individual in cluster *i* was selected for the survey
m = number of sample clusters selected
M_i = expected measure of population size for cluster *i*
M = total measure of size for the survey universe ($M = \Sigma M_i$ = sum of the population sizes of all clusters)
n_i = number of subpopulation members selected in cluster *i*
N_i = Total number of subpopulation members observed in cluster *i*.

An equal number of individuals is selected from each cluster at the second stage. Thus, M_i and N_i are equal and they cancel each other out. Since n_i is the same for each cluster, all clusters will have an equal sampling probability of n_i/M. As such, this design results in a self-weighted sample. Thus it is not necessary to apply sampling weights to obtain valid point estimates, but it is still necessary to account for the sampling design when calculating the standard errors, CIs and *P*-values. In addition, adjustment for other issues such as nonresponse is still needed. Should sampling frames be updated during the survey, confirm that M_i and N_i remain equal. If not, then consider another, more appropriate weighting approach.

Scenario 2:

- First stage: selection of clusters with equal probability
- Second stage: "take all" strategy. In cases where the number of eligible individuals in a venue is less than the target number of participants at each sampling event, all eligible individuals in the venue will be sampled.

$P_i = (m/M)$

where

P_i = probability that a subpopulation member in cluster *i* was chosen for the survey
m = number of sample clusters selected
M = total number of sample clusters in the sampling frame

Because all subpopulation members present on the randomly chosen day are selected for the sample, the second-stage sampling probability is equal to 1.0 and is not shown here. This design also results in a self-weighted sample, which means it is not necessary to apply sampling weights to obtain valid point estimates, but it is still necessary to account for the survey design when calculating the standard errors, CIs and *P*-values. In addition, adjustment for other issues (e.g. nonresponse) is still needed.

Scenario 3:

- First stage: selection of clusters with probability proportional to size (PPS)
- Second stage: subsampling

$P_i = (m * M_i/M) * (n_i/N_i)$

where

P_i = probability that an individual in cluster *i* was selected for the survey
m = number of sample clusters selected
M_i = expected measure of population size for cluster *i*
M = total measure of size for the survey universe ($M = \Sigma M_i$ = sum of the population sizes of all clusters)
n_i = number of subpopulation members selected in cluster *i*
N_i = total number of subpopulation members in the cluster *i*

This design results in a non-self-weighting sample; thus it is necessary to apply sampling weights during analysis.

Scenario 4:

- First stage: selection of clusters with equal probability
- Second stage: subsampling

$P_i = (m/M) * (n_i/N_i)$

where

P_i = probability that an individual in cluster *i* was selected for the survey
m = number of sample clusters selected
M_i = expected measure of population size for cluster *i*
M = total number of sample clusters in the sampling frame
n_i = number of subpopulation members selected in cluster *i*
N_i = total number of subpopulation members in the cluster *i*

This design results in a non-self-weighting sample; thus it is necessary to apply sampling weights during analysis.

Step 2: Create the sampling weights specific for the sampling design
After the appropriate sampling probability for the survey design has been calculated (Step 1), it is converted to a sampling weight as follows:

where

$$w_i = 1/P_i$$

w_i = sampling weight in the *i*th cluster

P_i = probability of selection in the *i*th cluster.

Weighting for respondent-driven sampling

Respondent-driven sampling (RDS) data are weighted using information from the participants' self-reported social network size (or degree) and recruitment patterns.

Because this weighting uses information from the recruitment process to calculate the sampling probabilities, weighted analysis for RDS differs from the analysis for other sampling methods in several important ways:

- it requires specialized software[2] to analyse the data.
- some estimators (e.g. RDS I and RDS II) take the RDS recruitment matrix into account; thus sampling probabilities (and, therefore, the weights) are generated for each variable and each record analysed by the software. For example, the sampling weight for a young female when analysing the "age" variable differs from the sampling weight when analysing the "sex" variable. The Gile SS estimator produces one weight that is applied to all variables for a single record. New estimators, with differing strengths and weaknesses, are being developed. Hence, the literature should be continually reviewed to find the appropriate estimator for each dataset.

Calculating response rates and weighting data

There are many ways to calculate response rates for CCS and TLS. The same method should be used for all sampling events in the survey. The method should be clearly explained in all reports or publications. Some methods are:

- number who take eligibility screener ÷ number approached;
- number who take eligibility screener ÷ (number approached × % eligible among those taking the eligibility screener);
- number who take survey ÷ number approached; and
- number who take survey ÷ (number approached × % eligible among those taking the eligibility screener).

Weighting in TLS is a complicated process:

1. The weight (1 ÷ probability) for a person at a given event is:

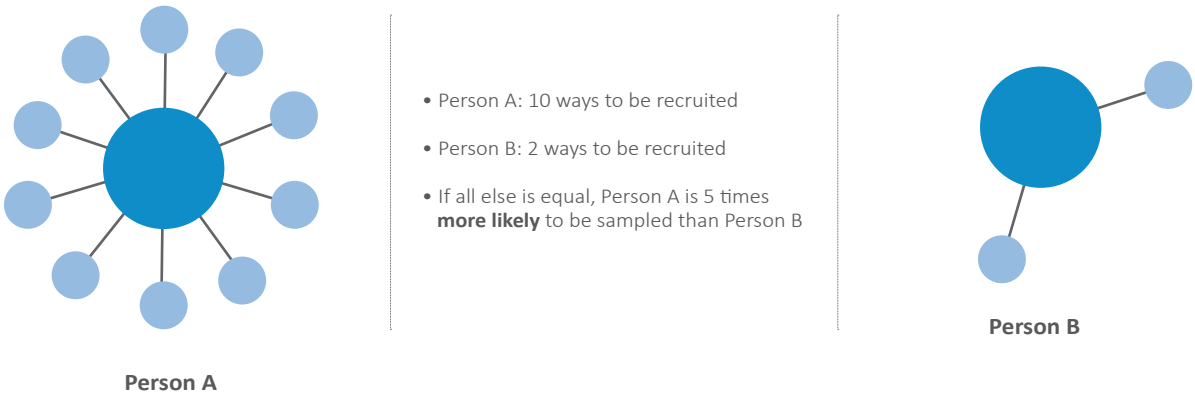
   [(pre-event count + entry count) × (eligible proportion among those taking screener)] ÷ (number interviewed)

2. Then, to generalize to venues, multiply those weights by:

   The number of eligible venue–day–times (VDTs) the event's venue has ÷ the number of events at which sampling occurred. A venue with many VDTs would receive more weight than a venue with few VDTs.

3. Modify each person's weight proportionally by their venue attendance frequency.

Figure C-1.1 Probability of selection



Source: Personal communication from Lisa Johnston, based on work done in 2012 at University of California, San Francisco

- there is no pre-existing information about the population (e.g. the sampling frames in cluster sampling) used in the weighted data analysis. Complex analytical techniques are used on the sample data to generate the sampling probabilities after the survey recruitment is finished.

Social-network size (degree)

The probability of selection is based on each participant's social-network size. This is measured by the number of people the participant knows who fulfill the eligibility criteria for the survey, and whom they have seen in a predefined period. Each participant's data are weighted by the inverse of the network size. Those with a small social-network size have a higher weight and those with a large social-network size have a lower weight.

The response to the question of network size acts as an estimate of multiplicity: it is a proxy for a measurement of the number of persons who might be able to recruit the participant. Individuals with larger network sizes are more likely to receive a coupon (see Figure C-1.1); hence, their data are given a lower weight. Individuals with a smaller network size are less likely to receive a coupon; hence, their data are given a higher weight.

Recruitment patterns

Participants might be more likely to recruit people similar to themselves, thus introducing additional differences in the chance of being recruited for the survey. The coupon system in RDS is designed for researchers to trace who recruited whom in an RDS survey. This information is used to create weights – referred to below as "recruitment weights"– that adjust for possible selection bias from recruitment preferences of individuals and different preferences among groups. Assess the sample for convergence and bottlenecks during implementation and after data collection has ended. The website of the Hard-to-Reach Population Methods Research Group[3] website or the WHO/UNAIDS guide to using RDS Analyst *(5)* can help users learn how to use the tools in this software that assess convergence in RDS Analyst.

1. Recruitment effectiveness. This component of the recruitment weight adjusts for the differences between groups in the probability that a person with a certain characteristic will recruit someone, regardless of the characteristic of the person they recruit. Groups with better recruitment effectiveness might end up being overrepresented in the sample *(6)*, which happens frequently when the population is similar or has high homophily (i.e. a tendency for individuals to know and socialize with people who have similar characteristics).
2. Differential recruitment between groups. This component of the recruitment weight adjusts for the probability that people with a certain characteristic will be recruited by someone with a certain characteristic, which may or may not be the same as that of the recruiter. The weighting attempts to account for the probability that someone will be recruited, based on the characteristic of the recruiter. These recruitment patterns are usually the result of homophily. For example, young people may have closer social-network ties among themselves than with older people, and therefore may differentially recruit other young people.

Homophily is a diagnostic statistic that describes the mixing patterns in networks and is calculated by RDS software. In RDSAT, homophily can be positive or negative, ranging from –1 to 1 *(7)*. In extreme cases, separate RDS samples within homophilous groups can be used *(8)*. Ideally, investigators should monitor homophily during data collection. If homophily is deemed

[2] For example, RDS-Analyst (http://hpmrg.org/software/) or RDSAT (http://www.respondentdrivensampling.org/).

[3] http://www.hpmrg.org/

high, investigators can improve recruitment training to encourage participants to recruit randomly from their networks, or can add new seeds. The homophily statistic also helps explain why the point estimate has a large CI if the homophily is high:

- homophily = 1: always recruits from own group;
- homophily >0 to 1: preferential recruitment of group with similar characteristics;
- homophily = 0: no preferential recruitment;
- homophily –1 to <0: preferential recruitment of groups with different characteristics; and
- homophily = –1: always recruits from other than own group.

Many analysts consider a value of between –0.3 and 0.3 as "normal" homophily. If homophily is outside this range (i.e. lower than –0.3 or higher than 0.3), differences between the unweighted sample data and weighted estimates may be more prominent.

In RDS Analyst, homophily hovers around 1, indicating no homophily. A value of 1.3 or greater is considered high homophily (personal communication with Mark Handcock, University of California at Los Angeles, 26 January 2016).

Differential recruitment among groups is usually limited to socially salient characteristics, that is, characteristics that can be observed in others and upon which people form friendships. For example, high homophily (above 0.3) is commonly observed in estimates of preferred drugs among people who use drugs, because they form friendships and drug-acquisition networks around the specific drugs they use. Conversely, homophily close to zero is observed in characteristics that are either not visible or not important for forming relationships, such as month of birth.

Weighting for recruitment patterns makes RDS analysis very different from analysis of data with other sampling strategies. Several estimators can be used to create RDS weights, with more being developed. Some estimators, such as Gile's Successive Sampling Estimator, create a single weight for each participant. Some, such as RDS I and RDS II, create a weight for each variable and each participant. Therefore, because the recruitment weights are based on the characteristics of the recruitments among groups, each record in a dataset will have a different weight for each of the variables analysed. For example, in a survey report with 500 respondents that has 100 indicators, 100 weight variables will be generated automatically by the specialized software (one for each indicator in each record).

C-1.3 Analytical methods for BBS data

Currently, BBS are used to observe associations between the outcomes and behaviours or characteristics in the survey population. BBS data may be analysed within a single survey round, or across three or more rounds in the form of trend analysis, if methods and sampling approaches are similar among those rounds. Ideally, RDS survey results should be compared with subsequent RDS survey results in the same population with the same tools (the same is true for TLS). To be comparable, surveys must have been performed in similar locations and populations. However, if the different survey techniques (RDS and TLS) have been used correctly on the same sampling population, it should still be possible to assess trends, because both will simulate simple random sampling (SRS). Whatever the technique used, trends must be handled and interpreted with caution. The analyses described here require statistical software packages that allow analysts to account for complex survey designs with components such as primary sampling units (PSUs), stratifying variables and weights to generate appropriate estimates. Most survey reports present results as weighted means for continuous variables (e.g. age) or weighted proportions for categorical variables (e.g. marital status), with visual analysis (graphics) of variable distributions or HIV-prevalence trends over time. Manuscripts might answer specific research questions that require more complicated statistical analyses (e.g. regression) that are beyond the scope of these guidelines.

In a survey report, the unweighted denominator, weighted mean or proportion, and CI (e.g. 95% CI) for each variable should be reported in summary tables (*9*). Table C-1.2 is an example of a summary table of sociodemographic and risk characteristics. The design effect, calculated from the survey data, can also be presented for the main variables. This enables readers to compare the effective sample size used to a simple random sample.

Table C-1.2 Example of a summary results table of sociodemographic and risk characteristics

Characteristic	n (unweighted)	Sample % (unweighted)	Estimated population % (weighted)	95% CI	Design effect
Age (years)					
15-19					
20-24					
25-29					
30-34					
35+					
Etc.					

CI, confidence interval; n, sample size

A note on denominators

Denominators must be well defined and adjusted when certain sequences of questions are asked, based on previous responses. In surveys, some questions are designed to prompt follow-up questions. These questions are referred to as "filter" questions and are often dichotomous (i.e. they have only two response values such as "Yes" or "No"). If a participant responds "Yes" to a filter question, this prompts one or more follow-up questions relevant only to those with "Yes" responses, acting as a type of filter. Conversely, a "No" response skips these follow-up questions because they are not appropriate for the participant based on previous responses, and the interviewer moves on to the next survey question. The term "skip pattern" describes this pattern of questions and responses, and provides information on denominators for each question. The resulting filter creates subsets of the data based on specific criteria (responses). Once applied, only participants who meet those criteria should respond. For filter questions, the denominator should be adjusted to reflect the participants and the criteria, depending on the intent of the analysis.

Generally, respondents with missing values are excluded from the denominator to compute percentages. For example, if one wants to estimate the proportion of female sex workers (FSW) who used a condom the last time they had anal sex, the analyst might want to keep in the denominator only those who reported ever having had anal sex.

However, in certain situations, it can be more meaningful to include participants with missing values in the denominator. For example, a survey might use the filter question "Have you ever been to school?" Participants who replied "No" are then excluded from the next question about the highest school level completed. When computing the proportion of respondents who completed secondary school or higher, the analyst should include in the denominator those who never went to school.

Biomarker data can also present denominator challenges. One such example is indeterminate HIV test results. Such results should be analysed as their own category (e.g. positive or reactive, negative or nonreactive, or indeterminate) and should be included in the denominator. Another analytical challenge is presented with viral load, a clinical measure of the number of viruses in a person's blood. Viral load is usually reported as copies of HIV/mL of blood, and an aggregate viral load (across all tested participants) can be calculated for the sample. However, when the viral load is below the threshold needed for detection by the specific test used, the result is classified as "undetectable" (e.g. <50 copies/mL with some current tests). Undetectable viral-load results need to be included in data analysis of viral load; the use of an imputed value representing the mid-point between zero and the threshold of detection is recommended (e.g. using 25 if the threshold of detection is 50 copies/mL).

Alternatively, and more conservatively, the threshold value itself may be used for imputation (50 copies/mL, in this example). Depending on the survey objectives, viral load can be reported as an arithmetic mean, a geometric mean (mean logarithm of the viral load), a median, or categorically with several epidemiologically or clinically meaningful categories. For more information on measuring and analysing viral load, see the 2017 Global AIDS Monitoring guidelines *(10)*.

C-1.3.1 Univariate analysis

Descriptive statistics summarize the sample, but do not test any hypotheses. They include measurements of centre such as means and medians, with corresponding measures of uncertainty (spread) around the centre. The *standard deviation* measures the precision of the mean, the *variance* measures how far a data point deviates from the mean, and the *range* represents the lowest value to highest value. The distribution of continuous survey variables, such as age or number of sex partners, can be reported as weighted means with 95% CI.

A proportion indicates the size of a part compared to the whole, with values falling between 0 and 1. Percentages, which are proportions multiplied by 100, are one way of expressing a proportion; for example, HIV prevalence. In survey reports, the distribution of categorical values are presented as weighted proportions with a 95% CI. The numerator is the count of participants with a specific characteristic (e.g. a positive HIV test result). The denominator includes all participants in whom the characteristic was measured (e.g. number tested for HIV). The counts in the numerator and denominator are multiplied by the appropriate weight (or weights).

C-1.3.2 Bivariate analysis

Bivariate analysis assesses the relationship between two variables. Typically, one of the variables is considered a "dependent" (outcome) variable (e.g. HIV status); the other is considered an "independent" (explanatory) variable (i.e. a characteristic, behaviour or exposure that explains the change in the dependent variable). Bivariate analyses of simple "yes/no" explanatory and outcome variables can be reported as weighted proportions, standard errors and 95% CIs. The Chi-square test measures the relationship between two categorical variables (e.g. a 2 × 2 contingency table). Statistical tests are accompanied by a *P*-value, which is the probability of an event given the null hypothesis is true. Chi-square tests the odds or risk of an event, given one or more independent variables; results are reported as crude (unadjusted) odds ratios or relative risk (measures of association between an exposure and an outcome), with associated 95% CIs. Logistic regression is another method of testing relationships between a dichotomous dependent and one or more independent variables while controlling (or adjusting) for potential confounding variables. Logistic regression also yields adjusted odds ratios with 95% CIs. Linear regression may be used with continuous outcomes.

Stratified analysis

Stratification is the classification of a survey population into subgroups or categories ("strata") on the basis of selected characteristics. For example, a survey population may be stratified by sex (male or female), age group (<24 years, 25–34 years or 35+ years) or residence (urban or rural). Stratified analysis may be used to examine a subpopulation of interest, or as a method of controlling for a particular characteristic that might be a confounder. A confounder is a variable that correlates with both the independent (explanatory) variable and the dependent (outcome) variable. When stratum-specific odds ratios are relatively similar, it is acceptable to pool results, but when the ratios differ, results must be presented separately.

Multivariable analysis

Often, it is necessary to simultaneously control (adjust) for multiple confounding variables. If stratified analysis were used for this purpose, some strata would contain very small numbers, thus reducing the power to detect any associations. Multivariable analysis, or statistical modelling, is a method used to control for multiple confounders simultaneously. Multivariable regression can be performed after exporting the RDS data to another software program; some programs (e.g. RDS Analyst and Stata), which can use RDS analytical tools, can directly perform the regression. Multiple logistic regression, a type of multivariable analysis used with dichotomous outcomes, is one of the most common statistical operations performed with BBS data to generate adjusted odds ratios (measures of association between outcome and multiple explanatory variables) with 95% CIs.

A consultation of statistical experts convened at the Centers for Disease Control and Prevention, Atlanta, Georgia, in 2012 recommended applying sampling weights that are created for RDS analysis during multivariable regression when analysing RDS data. Standard procedures for analysing sample survey data can be used to perform regression analysis of RDS data. The seed or shared recruiter should be specified as the primary sampling unit (i.e. cluster variable). Ideally, regression analysis should be performed using exponential-family random network models. These are computationally intense and not for the average user. The ERNM models can provide new approaches to network regression and multilevel modelling. Models include joint modelling of Y and X, exponential random graph models; Gibbs measure and conditional modelling of Y given X (personal communication from Berchenko Y, Frost S, Gile K, Handcock M, Heckathorn D, McFarland W, Salganik M, Szwarcwald C, Thompson S, White R on 22 August 2012).

Other analysis methods

More advanced investigations of repeated cross-sectional survey data include age–period–cohort analyses, which are based on the theory that cohort membership influences behaviour to a similar degree as other demographic factors. Other cohort or time-dependent statistical techniques that have been used to analyse repeated cross-sectional surveys include pooled logistic

regression, Cox proportional hazards regression, Kaplan-Meier curves, multilevel/hierarchical modeling, and generalized estimating equations to measure estimates over time.

These analyses are beyond the scope of these guidelines; seeking assistance from a statistician is recommended when investigating these complex analytical approaches.

C-1.4 Combining multiple surveys and subsetting data for analysis

Combining multiple surveys

In some situations, investigators need or want to combine data from multiple survey locations or periods to produce a single set of estimates. Combining data from multiple surveys should be performed with caution and only done when measures in each survey are comparable. For example, three cities in a country might have implemented surveys simultaneously and investigators want to combine survey data from each city for a national estimate. If combined without accounting for the unique characteristics (e.g. variance estimates) of the three surveys, the resulting estimates will be inaccurate. Thus, for each BBS sampling design, weighting needs to be adjusted to account for how the data will be combined (e.g. distinct geographical locations or periods have different variance estimations). Interpretation must consider that the final estimates are a combination of distinct samples with their own characteristics. Local data combined to produce national estimates need to be identified as such, rather than presented as nationally representative estimates.

Subpopulation estimates

There are situations that require estimates and inference for subpopulations. When performing a subpopulation analysis, the entire survey dataset should be retained to keep the integrity of the sampling-design information in order to correctly calculate the standard errors of the estimates. For example, if analysts want to report results for each step of the HIV-care cascade, they might select the subpopulation of survey participants with laboratory-confirmed HIV and then perform analyses on only those participants, rather than on all survey participants. Special commands for these subpopulation analyses are available in survey software packages and must be used on the entire survey dataset to maintain the integrity of weighting adjustments made on the *entire* survey dataset. Creating new datasets with only the subpopulation is not recommended because it will alter the weighting and may produce incorrect estimates, CIs and *P*-values.

C-1.5 Comparing surveys

Examining differences between two surveys

Investigators commonly want to determine whether there are statistically significant differences between estimates (either percentages or means) from two surveys. If the surveys have independently selected samples with similar populations, and if the methods and instruments are alike, results from the two surveys can be used to formally test for differences in estimates, under certain assumptions. This comparison is often performed using a Z-statistic directly calculated from the point estimates and the associated CIs.

To illustrate, consider the case in which we have conducted two RDS surveys at two different times, using the same methods and data collection instruments for the same populations. Data analysis of the first survey provides an estimate of HIV prevalence of p_1 with a $(1 - a)$% CI of (l_1, u_1) after accounting for the weights and sampling design. Similarly, data analysis of the second survey gives an estimate of HIV prevalence of p_2 with a CI of (l_2, u_2). If the CIs are roughly symmetric about the point estimates, then the standard errors (se) and variances (var) for the estimates are approximately:

$$\text{se}_1 = \frac{u_1 - l_1}{2 * Z_{1-a/2}} \quad \text{and } \text{var}_1 = \text{se}_1 * \text{se}_1$$

$$\text{se}_2 = \frac{u_2 - l_2}{2 * Z_{1-a/2}} \quad \text{and } \text{var}_2 = \text{se}_2 * \text{se}_2$$

where $Z_{1-a/2}$ is the $1 - a/2$ quantile for the standard normal distribution
(e.g. for 95% CI, $a = .05$ gives $Z_{.975} = 1.96$)

The Z statistic for the difference in the survey estimates $p_2 - p_1$ is calculated as:

$$Z = \frac{p_2 - p_1}{\sqrt{\text{var}_2 + \text{var}_1}}$$

The *p*-value for the Z-statistic is calculated using the quantile function of the standard normal distribution. If the $(1 - a)$% CIs of the point estimates of the first and the second surveys do not overlap, then the difference between the estimates is necessarily statistically significant, because the $(1 - a)$% CI about the difference between the two point estimates will not contain zero. However, even if the CIs overlap, the Z-statistic, as calculated above, may still be statistically significant.

Table C-1.3 Example of a table for an analysis of differences between two surveys

HIV Status	Survey A (2012) n=53	Survey B (2015) n=55
Positive	39	30
Negative	14	25
	p_1= 39/53 = 0.736	p_2=30/55 = 0.545

For example, let HIV status results in Survey A (2012) be 39 positive and 14 negative, and in Survey B (2015) be 30 positive and 25 negative (Table C-1.3). To test for a significant difference in HIV-status estimates between 2012 and 2015, the Z-statistic is equal to 2.0593, which gives a *p*-value of 0.0394. Small *p*-values (e.g. $p < 0.05$) indicate statistically significant differences between the surveys for the characteristics (point estimates).

If we are estimating means rather than proportions, the procedure is similar, except that it uses the estimated mean from the two surveys, m_1 and m_2, and their CIs instead.

Note that this procedure is only approximate and assumes that the CIs are symmetric about the point estimates and that the two surveys are independent (e.g. done at different times). If these assumptions cannot reasonably be made, then a statistician should be consulted about more appropriate methods for formally testing for significant differences.

Trend analysis

Trend analysis measures changes in weighted BBS estimates among three or more points (surveys) in time. Comparisons should be made only when measures and methods are the same in all surveys; best practice in BBS calls for repeated surveys using the same instruments in identical populations. Most trend analyses should be conducted by producing weighted estimates for each variable for each round of a survey, and then determining whether there is a difference between rounds (Table C-1.4). Visual trend analysis of comparable data (e.g. weighted mean age of participants or HIV-prevalence estimates) over time can be presented with 95% CIs (Figure C-1.2). Testing for statistical differences in comparable data over time can be performed using regression methods, with the guidance of a statistician. Combining results from surveys with different sampling strategies (e.g. RDS and TLS) or population definitions (e.g. SW and transactional sex) should be avoided.

Table C-1.4 Example of a table shell for an analysis of differences between two surveys

Characteristic	Survey A (2012) N =		Survey B (2015) N =		Difference (*p*-value)
	n	% (95% CI)	n	% (95% CI)	
HIV status					
Positive	# positive	% positive of all tested, p_1 (95% CI)	# positive	% positive of all tested, p_2 (95% CI)	
Negative	# negative	% negative of all tested (95% CI)	# negative	% negative of all tested (95% CI)	

Figure C-1.4 Example of visual trend analysis of HIV prevalence with 95% confidence intervals from four rounds of BBS in the same population at Location X, 2005–2013

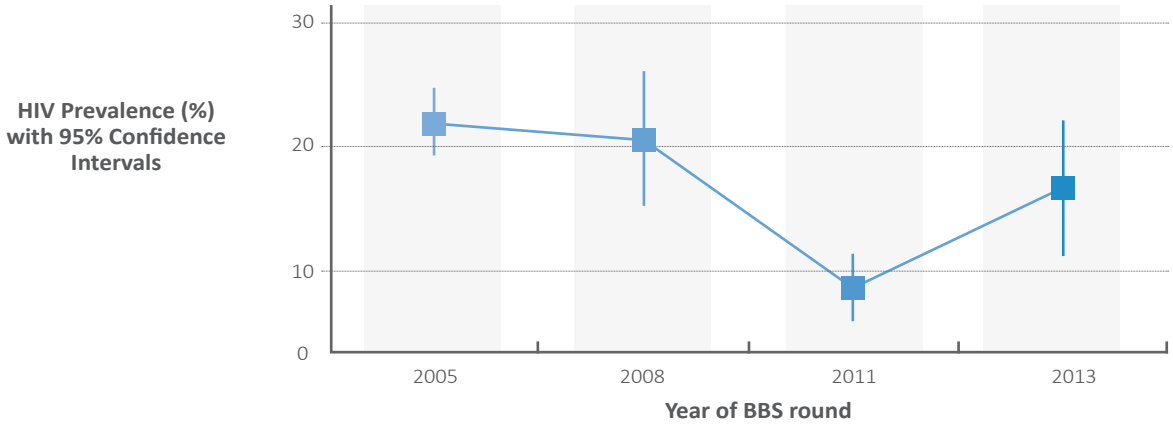


Table C-1.5 provides a comparative overview of the analytical approaches for BBS discussed above.

Table C-1.5 Overview of common analytical approaches for BBS

Approach	Description	Purpose	Example
Univariate analysis	Frequency of a variable value or indicator relative to one denominator	To describe the distribution of a characteristic, behaviour or outcome in a population	12.5% of 800 MSM surveyed were HIV+; the mean age of sex workers surveyed was 26 years
Bivariate analysis	Frequency of a variable value or indicator relative to two denominators (e.g. condom use at last sex among HIV+ MSM and HIV– MSM)	To compare the distribution of a characteristic, behaviour or outcome in two groups; or to measure the association between two variables	Frequency of HIV infection was similar (P >0.05) among 200 women (18.3%) and 350 men (17.9%) in the sample; HIV+ MSM were 2.3 times more likely to report inconsistent condom use than were HIV–MSM (OR 2.3; 95% CI:1.9, 2.6)
Stratified analysis	Bivariate analysis for two or more strata or categories of a third variable (e.g. a subpopulation such as age group, marital status, or urban or rural residence); allows assessment of the measure of association in different strata.	To compare distribution of a characteristic, behaviour or outcome in two groups by multiple strata of a third variable (potential confounder)	Among MSM aged <25 years, the odds of being HIV+ were 2.1 times higher in those reporting inconsistent condom use than in those reporting consistent condom use; whereas among MSM aged >25 years, the odds of being HIV+ were 4.2 times higher in those reporting inconsistent condom use
Multivariate analysis	With dichotomous outcomes, use logistic regression to analyse potential association with multiple variables and indicators	To assess independent associations between outcome (dependent) and exposure (independent) variables	After adjusting for age, marital status and urban/rural residence, HIV+ MSM were 5 times more likely to have an STI than were HIV–MSM
Comparing surveys	Comparison of a characteristic, behaviour or outcome variable or indicator over time or between two surveys	To assess whether significant differences exist in outcomes or characteristics over two or more BBS rounds	The HIV prevalence among MSM in City X increased significantly over BBS rounds 1, 2 and 3
Trend analysis	Changes in same characteristic, behaviour or outcome variable or indicator over time	To assess whether significant differences exist in outcomes or characteristics over three or more BBS rounds	The HIV prevalence among MSM in City X increased significantly over BBS rounds 1, 2 and 3

OR: odds ratio; STI, sexually transmitted infection.

C-1.6 Suggested resources

University of California, Los Angeles (UCLA), UCLA stats: http://www.ats.ucla.edu/stat/

RDSAT 7.1 user manual: www.respondentdrivensampling.org

Harvard summary of survey analysis packages: http://www.hcp.med.harvard.edu/statistics/survey-soft/

C-1.7 References

1 Global Health Sciences. Toolbox for conducting integrated HIV bio-behavioral surveillance (IBBS) in key populations. San Francisco: University of California; 2016 (http://globalhealthsciences.ucsf.edu/pphg/gsi/epidemiologic-surveillance/ibbs-toolbox, accessed 30 August 2016).

2 CDC. National HIV behavioral surveillance system: men who have sex with men – Round 4: operations manual. Atlanta, GA: Centers for Disease Control and Prevention (CDC); 2014 (http://www.cdc.gov/hiv/pdf/statistics/systems/nhbs/nhbs-msm4-operations-manual--version-date-05-16-2014.pdf, accessed 4 August 2016).

3 Gardner MJ, Altman DG. Statistics with confidence. London, BMJ. 1994.

4 Aiken LS, West SG. Multiple regression: testing and interpreting interactions. Sage Publications. 1991.

5 WHO/UNAIDS. Introduction to HIV/AIDS and sexually transmitted infection surveillance: module 4 supplement. A guide to using RDS Analyst and NetDraw. World Health Organization (WHO)/Joint United Nations Programme on HIV/AIDS (UNAIDS);(http://applications.emro.who.int/dsaf/EMRPUB_2014_EN_1686.pdf, accessed 5 November 2016).

6 Erickson BH. Some problems of inference from chain data. Sociol Methodol. 1979;10:276–302.

7 McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. Annu Rev Sociol. 2001;27(1):415–444.

8 Wejnert C, Pham H, Krishna N, Le B, DiNenno E. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. AIDS Behav. 2012;16(4):797–806 (http://www.ncbi.nlm.nih.gov/pubmed/22350828, accessed 8 August 2016).

9 White RG, Hakim AJ, Salganik MJ, Spiller MW, Johnston LG, Kerr L et al. Strengthening the reporting of observational studies in epidemiology for respondent-driven sampling studies: "STROBE-RDS" statement. J Clin Epidemiol. 2015;68(12):1463–1471.

10 UNAIDS/UNICEF/WHO Global AIDS Monitoring 2017: Indicators for monitoring the 2016 United Nations Political Declaration on HIV and AIDS. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS)/United Nations Children's Fund (UNICEF)/World Health Organization (WHO); 2016. (http://www.unaids.org/sites/default/files/media_asset/2017-Global-AIDS-Monitoring_en.pdf, accessed on 31 August 2017)

11 Heeringa SG, Brady TW, Berglund PA. Applied survey data analysis. Boca Raton, CRC Press. 2010.

12 Korn EL, Graubard BI. Analysis of health surveys. New York, Wiley. 1999.

2. Use and dissemination of survey findings

Biobehavioural surveys (BBS) are only useful if their findings are used for action. BBS findings have many uses, ranging from programme planning and advocacy to guiding the development of future surveys. This chapter describes how to disseminate and use BBS data, including potential uses for BBS data, identifying a target audience, tailoring a message and determining a dissemination strategy.

Investigators should decide early (during the planning phase) how the survey data will be used. Unless data use is planned before data collection begins, some information that might be of interest to decision-makers may not get collected.

C-2.1 Potential uses for BBS data

Planning HIV programmes

BBS using probability-based samples provide population-level estimates of target population demographics, risk behaviours, burden of disease, and service uptake to inform HIV programmes. BBS data may be used to estimate the proportion of the target population that has been exposed to a particular intervention or service. Survey data can also play a role in estimating programme access, coverage and acceptability (for suggested data measures, see the questionnaire module on services uptake and indicators in Sections II and III of the Supplemental Materials). Proven interventions that have been underused may be expanded. Measures of interest that appear unchanged despite efforts may need a new approach, perhaps one that pays more attention to the cultural or social context that determines why people behave in a particular way. Additionally, BBS data can be used to identify areas where subpopulations need more support or to identify successes in working with a population.

Examples of the use of BBS data

Example 1: a survey among people who inject drugs (PWID) may estimate the proportion who were tested for HIV in the past 12 months and, among those who tested HIV positive, the proportion enrolled into care. The survey data may inform investigators on factors associated with PWID who are HIV-positive enrolling into care (e.g. through outreach testing services or the availability of care services tailored to PWID). These findings may indicate that expanding such services could increase the proportion of these people enrolling into care.

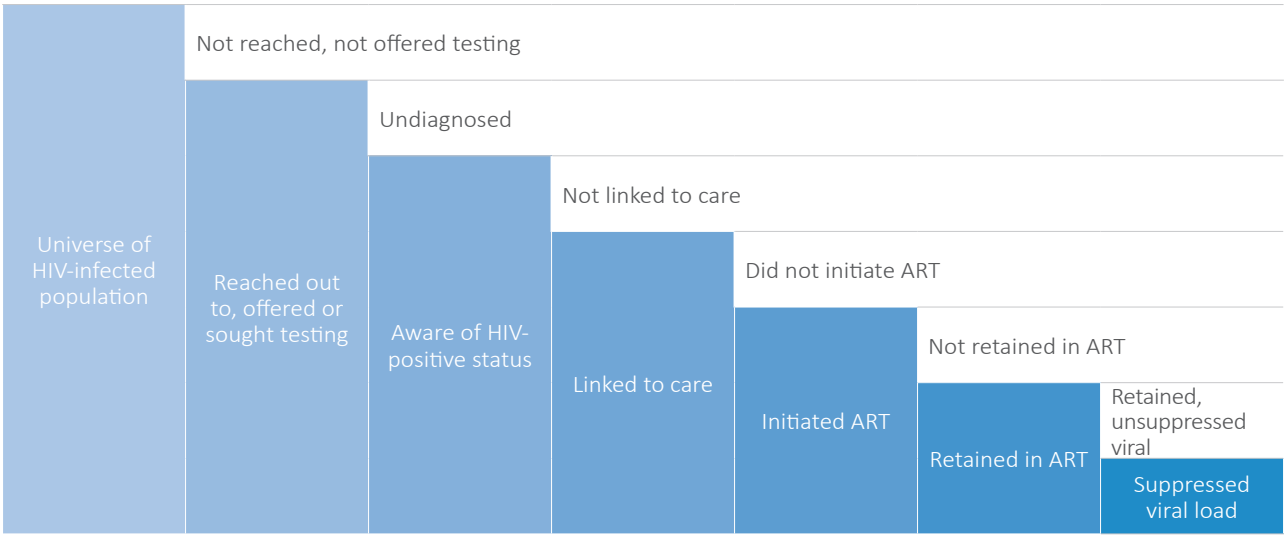
Example 2: a survey may find that sex workers (SW) who received peer-to-peer training are less likely to be victims of client-initiated violence than SW who did not receive the training. This finding could support expansion and increased funding of community-based training programmes. Various other studies have examined how surveys inform public-health action *(1-3)*.

Data from surveys with sufficiently large sample sizes (and sufficient numbers of respondents who are HIV-positive) may be of use in estimating the uptake of services along the continuum of HIV services, also known as the HIV prevention cascade and the "HIV cascade of care" *(2, 4)*. The cascade displays the proportions (or percentages) of the participants who have been exposed to or have taken up certain HIV-related services, including having been offered HIV testing, enrolling in care, being in treatment, and having a suppressed viral load (Figure C-2.1). Surveys with probability samples may even enable estimation of such uptakes on a population level. Further, because population-based surveys sample individuals who do and do not use services, these data can be used to characterize those who are outside the continuum of care. UNAIDS has a goal of ensuring that, by 2020, 90% of people who are HIV-positive know their status, 90% of all people diagnosed with HIV are receiving sustained antiretroviral therapy, and 90% of all people on treatment have suppressed viral load. High-quality data on these indicators for key populations will help target efforts to ensure the UNAIDS goal is met.

Advocacy

Survey findings can support advocacy for increased funding, programmatic expansion and legal protection, and highlight issues related to stigma and violence. Community-based organizations are often particularly effective users of survey findings. Helping organizations to understand survey findings enables them to take ownership of the data, and to use them in workshops or funding applications.

Figure C-2.1 Schematic of a continuum of HIV care and treatment services populated through survey data



ART, antiretroviral therapy

Presented appropriately to donors and international agencies, survey findings often play an important role in advocating for increased resources for activities that are not being adequately covered in government spending plans. The government health leadership or other government entities may implement policy changes based on new survey findings.

Inform future surveys

Previously completed surveys offer a wealth of practical experience that is invaluable for planning the next survey. For example, planners can learn about the sampling efficiency of the chosen design, whether the level of compensation was appropriate, and the usefulness of the collected data and laboratory measures.

C-2.2 Identifying a target audience

Target audiences vary in their scientific (or technical) literacy and interest in survey findings, and in their use of those findings. Therefore, the content of the message should be tailored to the intended audience. Some questions to consider before tailoring a message are:

- who is the target audience?
- what does the target audience already know about the subject?
- what does the target audience need to know?
- which findings are the most interesting or important for this audience?
- what are the specific messages for the target audience?
- what questions will the target audience have?
- how will this audience use the information from the survey?
- what is the best way to present and discuss the information with this audience?

Target audiences for BBS vary widely but should always start with the people from whom the data were collected: the survey's target population. Other audiences include programme planners and policy-makers, service providers, donors, nongovernmental organizations (NGOs) and community-based organizations, as well as the wider scientific community.

Target population of the survey

Representatives of the surveyed population should be among the first to be informed about the survey's findings in order to facilitate community-based organization work and advocacy, and to enable members to make informed choices about community-level efforts for HIV prevention. In addition, the target group might help the researchers interpret results and help plan the dissemination to other stakeholders. They may also have insight on how other stakeholders will use the results. When disseminating the survey's findings, the language used should be appropriate and without epidemiological jargon.

Service providers

Surveys can inform service providers about their overall reach and their reach concerning particular subgroups, and how they might improve access, acceptability and quality of services. Surveys may also inform service providers about the characteristics of populations whom they are failing to reach.

Policy-makers

Policy-makers have many pressing and competing priorities, and investigators who inform them should adjust the focus of the presentations to the needs of and level of detail required for the specific policy-makers.

Sometimes policy-makers will consider issues such as funding priorities, the enforcement of laws on the ground and treatment eligibility based on the behaviour of at-risk groups. Policy-makers also need information on the absolute burden of disease and the need for services (e.g. population size of target population and estimated number requiring services).

Scientific community

Survey findings published in peer-reviewed literature have the widest reach among the scientific community. Publications reviewing survey findings on a regional or global level often include only surveys published in journals indexed in databases (e.g. PubMed and Medline). Hence, any survey of acceptable quality should be published to ensure its lasting place in the scientific literature. Technical reports targeting a scientifically trained audience should also focus on methodological issues such as sampling design and statistical tests used.

Media and the public

Journalists from newspapers, radio and, occasionally, television can be important for communicating a survey's main findings. Investigators of surveys among vulnerable populations facing little or no stigma (e.g. students and transport workers) may choose to "go public" without putting the surveyed population at risk; however, stigmatized key populations should be protected from adverse media attention. For surveys of stigmatized populations, investigators should consult first with community-based organizations about the risks and benefits of briefing the media. When briefing, they should use simple language and focus on the survey's findings, their interpretation and recommendations, and should avoid any messages that could further stigmatize the survey's target population.

C-2.3 Tailoring a message

The careful selection of relevant data and its presentation in the appropriate way are fundamental to the audience's acceptance and use of the findings. Different audiences are at different "stages of change" according to some health-behaviour models that encourage public-health workers to identify the attitudes and knowledge of the audience, and to tailor the message appropriately. If the audience is unknowledgeable about the target population, but is receptive to the idea of assisting the population, then more data on the situation and how to help them would be appropriate. If the audience is suspicious or hostile to the population, then care is needed in deciding what information to provide. In the latter case, it may be appropriate to focus on a message that humanizes the key population, or the benefit to the general population of helping this key population.

The following are examples of target audiences and targeted information:

- a ministry of health will need to know which subpopulations are at greatest risk for acquiring HIV, and the population proportions undiagnosed and not enrolled into care and treatment, in order to prioritize the concept of "treatment as prevention".
- a community-based organization for SW will want to know about violence, stigma and access to services.
- donors will be interested in how uptake of services and treatment coverage for a particular key population compare to those in the general population in the same locale.

As a specific example, findings from an MSM survey that might be shared with policy-makers include the proportion of MSM who are HIV positive and who are on treatment, and the proportion of MSM who are bisexual and having sex with both men and women. These data might be compared to treatment uptake in the general HIV-positive population; the data may show that MSM are less likely to be on treatment and that these MSM have female sex partners who are therefore at increased risk. Together, these data can make a compelling case to policy-makers to see the HIV epidemic among MSM not in isolation but as part of a larger social and sexual network that facilitates HIV transmission among both MSM and the general population. Officials can draw conclusions from the data presented and suggest specific action to improve policies for HIV prevention and care.

C-2.4 Determining a dissemination strategy

After the target audience and message have been determined, the next step is to decide how to deliver the message. As discussed below, dissemination methods include:

- technical reports
- indicator reporting
- press releases
- issues briefs and fact sheets
- presentation slide sets
- manuscripts
- information cards for distribution to members of the population

Technical reports

A technical report is the most common information product of a BBS, but is rarely indexed in searchable databases. The report serves as a reference in subsequent publications and ensures that a record of the survey is publicly available. A technical report contains a complete description of the survey: the

formative assessment findings; recruitment methods and survey domains; data-analysis approach; and findings, including a presentation as data tables and graphs. The report should also provide contact information for the authors and the owners of the final dataset should additional analysis be necessary. Technical reports usually present relatively simple descriptive statistics, including numerators, denominators and the prevalence and distribution of specific characteristics. They can provide bivariate analyses of the primary outcomes. A report should focus on indicators that track risk, prevention efforts, or both, and include recommendations for policy and service delivery. The more user friendly the presentation of the data, the more likely it will be used. The technical report should be released at a workshop to ensure stakeholders are aware of the data. See Appendix I-34 for a suggested table of contents for a hypothetical survey report.

Indicator reporting

Survey results provide the data for standardized reporting to national governments, United Nations (UN) agencies and donors. Checking during survey preparation whether the data instruments will provide the appropriate data for reporting is good practice. The standardized questionnaires provided in these guidelines include questions that meet reporting requirements for UNAIDS, the United States President's Emergency Plan for AIDS Relief (PEPFAR), and the Global Fund to Fight AIDS, Tuberculosis and Malaria (the Global Fund). Section 3 of this document provides a list of common indicators.

Press releases and interviews

The mainstream media should be informed about a survey only if investigators can be assured that the safety of the population surveyed is not compromised. Ideally, investigators should discuss with representatives of the target population the benefits and risks of informing the public about the survey's findings. If informing the public is in the interest of the target population, consider a press conference or a workshop aimed at journalists, who appreciate a prepared short summary with take-away messages that facilitate their work and improve the accuracy of their reporting. Such press releases should contain minimal information on the methods used in the survey and focus on the findings, what they mean, and what to do about them. For interviews, predetermined communication objectives and prepared statements should be used. Whatever the question, conveying the main points in interviews is vital. See Appendix I-35 for a sample press release.

Issues briefs and fact sheets

The minister of health may not have time to read a 50-page report about HIV-related behaviour that includes information on sampling methodology and statistical tests, but may well be interested in learning whether the country's overall strategic plan for AIDS control is working in a particular high-risk group. The issues brief may highlight gaps in service provision, or successes in the uptake of particular programmes such as condom use or HIV testing. Such information should be packaged with information from other sources into a one-page or two-page brief that makes a compelling case for more HIV-prevention activities in a particular at-risk population.

Slide presentations

Slide presentations are helpful for presenting information in person to a group of people.

Manuscripts

Manuscripts published in peer-reviewed journals target the scientific community and create a lasting record of the survey. Most journals are indexed; therefore, readers can easily find the manuscript through scientific search engines. Publishing important findings in a peer-reviewed journal contributes to the wider knowledge base about the target population and the state of the HIV epidemic.

Presenting to challenging audiences

One type of audience deserves special attention: any group that has demonstrated the power to obstruct effective HIV-prevention efforts. Information prepared for these groups should take into account the concerns that form the basis of their opposition to prevention activities in the target population. For example, if religious leaders or politicians are opposed to HIV services for SW (e.g. condom distribution or training in negotiating safe sex), bombarding them with data showing that HIV prevalence is high or that poverty is often the root cause for taking up sex work is unhelpful. These data must be sensitively and strategically presented; for example, by showing that HIV services in this population do not promote sex work, and that prevention services benefit the families of SW and the wider population by lowering the risk of HIV transmission between SW and clients and their families.

See Appendix I-36 for more information on data presentation for clarity and impact.

C-2.5 Additional considerations for BBS data interpretation

Cause and effect

Inferences (interpretations) drawn from BBS data should be made in consultation with technical staff (e.g. statisticians and epidemiologists) to ensure that data are interpreted correctly. Demonstrating that behaviours or HIV prevalence change after programme implementation is one of the most effective ways to increase support for prevention and other activities. However, BBS are cross-sectional surveys. Data from a series of BBS over

time cannot be used to infer cause and effect in the effectiveness of interventions, only to present ecological[2] observations.[1]

One example is a change in HIV prevalence; such a change over two or more BBS rounds may or may not indicate a change in HIV incidence. HIV prevalence is a function of HIV incidence and mortality. Increased uptake of antiretroviral therapy (ART) should result in decreased mortality. Thus, an increase in people who are HIV-positive and are receiving ART may result in an increased prevalence despite a decline in incidence. In addition, numbers in some high-risk groups may fluctuate markedly; for example, in SW numbers through a rapid turnover of women or men in the industry or migration into or out of a city. Such changes can lead to fluctuations in the observed HIV prevalence. Changes in the prevalence of some sexually transmitted infections (STIs) (e.g. syphilis, gonorrhoea and chlamydia) may reflect changes in both risk behaviours (unprotected sex) and the uptake of treatment services. Correlating condom use with HIV prevalence is problematic, because condom-use behaviour is probed over a short period whereas HIV infection is a marker of chronic disease. Moreover, people aware of their HIV-positive status may be more likely to use condoms than those who are HIV-negative.

Presenting the estimated size of the key population

Most, or all, methods of estimating population size are based on assumptions that are hard to meet and difficult to assess. Therefore, the validity of these estimates is often unknown. When presenting size estimates, it is best to always present the accompanying confidence limits or uncertainty bounds, to present the results of any sensitivity analysis, and to discuss the possible limitations and biases, comparing them to previous or regional estimates. The estimates of population size should be rounded, because exact numbers imply a level of accuracy that cannot be achieved.

Timeliness

Survey data and findings lose utility over time; therefore, BBS findings should be presented and published as soon as possible. Guidelines for the reporting of data from surveys using respondent-driven sampling form a useful starting point for planning reports and publications *(5)*. Delays in disseminating survey findings can lead to delays in programme improvement, expansion or re-direction. Early dissemination may entail releasing results for key indicators through, for example, a presentation within 1–2 months of data collection, while report-writing is occurring. However, investigators should ensure that the preliminary findings will remain valid and will not change subsequently. Information may be discounted by an audience if it appears out of date or contradicts earlier reports. A full report can be distributed after data analysis and write-up is complete.

C-2.6 Relevant websites

Research utilization toolkit: https://www.k4health.org/toolkits/research-utilization

Eight strategies for research to practice: http://www.fhi360.org/resource/eight-strategies-research-practice

C-2.7 References

1 Emmanuel F, Adrien A, Athar U, Imran M, Reza T, Blanchard J. Using surveillance data for action: lessons learnt from the second generation HIV/AIDS surveillance project in Pakistan. East Mediterr Health J. 2011;17(8):712–718 (http://www.ncbi.nlm.nih.gov/pubmed/21977576, accessed 4 August 2016).

2 Hladik W, Benech I, Bateganya M, Hakim AJ. The utility of population-based surveys to describe the continuum of HIV services for key and general populations. Int J STD AIDS. 2016;27(1):5–12.

3 Sweeney P, Gardner LI, Buchacz K, Garland PM, Mugavero MJ, Bosshart JT et al. Shifting the paradigm: using HIV surveillance data as a foundation for improving HIV care and preventing HIV infection. Milbank Q. 2013;91(3):558–603 (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3790525, accessed 4 August 2016).

4 Kilmarx PH, Mutasa-Apollo T. Patching a leaky pipe: the cascade of HIV care. Curr Opin HIV AIDS. 2013;8(1):59–64.

5 White RG, Hakim AJ, Salganik MJ, Spiller MW, Johnston LG, Kerr L et al. Strengthening the reporting of observational studies in epidemiology for respondent-driven sampling studies: "STROBE-RDS" statement. J Clin Epidemiol. 2015;68(12):1463–1471.

[1] Ecological studies are those in which the unit of observation is the population rather than the individual.

Section C

3. Evaluating survey implementation

This chapter discusses survey evaluation, which is conducted after a survey has been implemented. Evaluation examines broad, overarching areas such as how well the survey was implemented, adherence to the protocol and standard operating procedures (SOPs), and the quality of survey and laboratory data.

Evaluation seeks to identify lessons that, if learned, can improve the next survey. In contrast, monitoring is the ongoing assessment of survey implementation, including sampling, recruitment and consent of survey participants, administration of data-collection instruments, data quality, specimen collection and processing, and referrals. Thus monitoring is conducted during a survey and results in immediate corrective action. Data from monitoring activities can also be used for evaluation.

There are many different types of evaluation. To provide details on specific evaluation methodologies is beyond the scope of these guidelines. Table C-3.1 lists general evaluation topics that are useful for biobehavioural surveys (BBS). Monitoring should have determined whether the concerns and questions listed occurred during a survey, whereas evaluation should determine how these items were or were not detected, their cause, and how they were addressed.

C-3.1 Data collection for evaluation

Evaluation can be conducted internally (by the investigators) or externally by a third party. Data for evaluation can be generated qualitatively (e.g. obtaining feedback from field staff and stakeholders, or reviewing participant feedback) or quantitatively (e.g. examining data quality or reviewing test results).

Table C-3.1 Evaluation topics and related questions to consider for biobehavioural surveys

Topic	Specific concerns and questions
Sampling	Sampling conducted according to protocol and SOPs
Recruitment and enrolment	Duplicate enrollees People posing as members of target population Fake or invalid coupons, coupon uptake (RDS) Inappropriate compensation, selling of coupons (RDS) Any respondents enrolled who were ineligible but received coupons (RDS)
Data collection	Missing interviews Partially completed interviews Identifying and rephrasing questions that: • were difficult to understand • led to refusals • led to extreme values
Laboratory	Proportion of test results returned Proportion of test results delayed Missing and poor-quality specimens Poor-quality testing Insufficient supply of test kits Testing errors

Topic	Specific concerns and questions
Human subjects	Breaches in confidentiality Informed consent procedures followed
Participant burden and acceptance	Time spent for survey participation Aborted interviews or survey procedures
Data-collection costs	Actual vs budgeted data-collection costs
Biomarker measurement	Inclusion of external quality assessment to validate results (see Chapter A-6)

SOPs, standard operating procedures; RDS, respondent-driven sampling

Obtaining feedback from field staff and stakeholders

Field staff are a valuable source of information because of their direct interaction with participants and their close involvement in survey activities. Together with the investigators, field staff can share what worked well, what challenges arose, and how they overcame those challenges. In addition to guiding planning for future surveys, this information can provide context for interpreting results.

Stakeholders (e.g. ministry of health, donors and community-based organizations) should be asked whether the survey findings met their needs, and what additional data they want collected in future surveys.

Reviewing participant feedback

Participant satisfaction is important for a successful survey. Unhappy participants may discourage others from joining; thus, participant feedback should be monitored during survey implementation, reviewed again after implementation and synthesized with other data from other sources. Feedback may be gathered after participants complete the survey, or through interviews, a suggestion box or email. If such interviews are systematic, the procedure should be described in the survey protocol.

Data-quality evaluation

Data quality refers to the completeness, validity, consistency and accuracy of data. Data-quality assurance (e.g. checking for missing values, extreme values and logic errors) should form part of ongoing monitoring activities. Evaluation of data quality looks at *overall* data-quality issues. For example, if an evaluation finds considerable missing data, the reasons for these omissions must be determined (e.g. data-entry error, incorrect skip patterns, data manipulation or analysis error, or corrupt files), and how to avoid similar future occurrences. Poor data quality is a serious issue that may result in missing data or incorrect estimates of the variables of interest. For guidance on assessing data quality, see Chapter B-2.

Review of specimen test results

At the end of a survey, investigators should determine the proportion of test results returned on time or delayed, the number of missing specimens, and the number of specimens in storage, if applicable. The evaluation should also examine the quality and quantity of specimens, the quality of test procedures, and the reporting of correct and incorrect test results. In addition, the evaluation should assess when and how problems were identified during the testing process, what corrective action was taken, when the action was taken, and the steps required to prevent similar problems in future surveys.

C-3.2 Using evaluation findings

An evaluation report should be developed after the evaluation is complete. Findings from the evaluation should be shared with field-based and non-field-based survey staff. Survey staff must use the evaluation findings for action. Creating a table similar to Table C-3.1 may be useful, with an additional column on the right entitled “Recommendations for future surveys” that addresses each concern or problem identified during the evaluation.

Section C

4. Data sharing for public use

This chapter discusses the rationale for and means of providing third-party researchers access to individual-level data from biobehavioural surveys (BBS). Most BBS are publicly funded and data are collected from subpopulations of the public. Thus, the public has a right to be informed of the survey findings and, with certain safeguards, to have access to and use of the data. Sharing anonymized individual-level survey data free of charge is good public-health practice because it maximizes a survey's utility. Data sharing allows other researchers to use the data in ways the original investigators may not have considered.

C-4.1 Data Sharing

Sometimes a survey's funding organization requires that survey data are made available to third parties. Examples of data sharing are rare for previous BBS, but best practices can be found in other survey areas. For example, data from the Demographic and Health Survey are routinely made available for interactive analysis or downloading from a website, because these data are from the US-based National Health and Nutrition Examination Survey (NHANES).[1 2]

Considerations when providing access to data

Breaches of confidentiality and misuse of survey data may harm the target population and must be prevented. The ultimate responsibility for securing data and preventing misuse lies with the survey's principal investigators. Individual-level survey data should be shared only if the risk of harm to the surveyed population can be minimized. This may require the vetting of potential data recipients by community-based organizations, and grouping data into larger categories to limit the potential for identification of participants. For example, participant age can be grouped into categories of 5–10 years.

Survey protocols should indicate if survey data will be made available to other researchers. Regulations concerning human subjects may require the inclusion of similar information in the consent form. Investigators should determine whether the investigators' or funders' institutional regulations, or national policies, mandate, recommend or caution against data sharing. Knowledge of national data-confidentiality regulations and standards should guide the decision to make data accessible to others *(1)*. Data-use agreements should describe the requirements for access to and use of data.

Preparing datasets for public use

Investigators should clean raw data to make them error-free before sharing, or provide the raw data, cleaning code, and resulting clean data file, in order to increase the transparency of data cleaning and management. Secondary variables may be added to facilitate analysis. When making survey data accessible, reference documents should be included for context, including the data dictionary, the data instruments, the survey protocol, guidance for appropriate attribution, and investigator contact information. A data dictionary explains the response values and shows the variable labels; the data instruments show the full question-and-response phrasing, as well as skip patterns; the survey protocol provides information about sampling design, eligibility criteria and other important design elements. Public-use datasets must be stripped of personally identifiable information; for example, dates of birth (day and month), photographic images, email addresses, phone numbers, street addresses, and social security or other identification numbers (IDs). Geographical information such as place of residence can be merged into larger area groupings or removed altogether.

Reporting standards to guide third-party researchers may also be outlined in order to protect participant identity; for example, not presenting values for less than five participants in a table cell.

[1] http://dhsprogram.com/Data/
[2] http://www.cdc.gov/nchs/nhanes.htm

Timing of data sharing

Survey data should be made available as soon as possible. Some agencies may have a policy of releasing data a specific number of months after survey completion.

Websites dedicated to data sharing

Websites dedicated to hosting and sharing research and survey data enable third-party researchers to access data. Examples of such websites include the Princeton University Office of Population Research (OPR) Data Archive,[3] Harvard University's Dataverse Network,[4] and the Inter-university Consortium for Political and Social Research at the University of Michigan.[5] These websites include information about policies for accessing and using datasets, including the format in which the data should be made available, and which metadata and information should accompany them. Third-party researchers can then access the data by abiding by the data-sharing website's rules and regulations. In most cases, a user must sign a data-use agreement in order to access the data; for some datasets, the user implicitly consents to the data-use agreement through the act of downloading the datasets. See Appendix I-37 for a sample data-use agreement.

C-4.2 References

1 Bernstein AB, Sweeney MH. Public health surveillance data: legal, policy, ethical, regulatory, and practical issues. MMWR Surveill Summ. 2012;61:30–34.

[3] http://opr.princeton.edu/archive/

[4] http://thedata.harvard.edu/dvn/

[5] http://www.icpsr.umich.edu/icpsrweb/ICPSR/

ISBN: 978-92-4-151301-2