

UNAIDS 2021

REFERENCE

PrEP target-setting for key and high-priority populations

Using the tools

Contents

- 2** Abbreviations
- 3** Entering data in the Estimates Summary
- 6** Example 1: Men who have sex with men—exposures approach
- 19** Example 2: Men who have sex with men—risk factors approach
- 23** Example 3: Female sex workers—exposures approach
- 35** Example 4: Transgender women—risk factors approach
- 43** Example 5: People who inject drugs—risk factors approach
- 50** Example 6: Adolescent girls and young women—risk factors approach

Abbreviations

BBS	biobehavioural surveillance
CI	confidence interval
DHS	demographic and health survey
HIV	human immunodeficiency virus
PHIA	population-based HIV impact assessment
PrEP	pre-exposure prophylaxis
PSE	population size estimate
RDS	respondent-driven sampling
UNAIDS	United Nations Joint Programme for HIV/AIDS
WHO	World Health Organization

Entering data in the Estimates Summary

We recommend reading through one or more of the fully worked examples in this document before using the tools. Instructions are also available in the individual worksheets.

The Estimates Summary sheet is the first worksheet in the tools. It is used to calculate the targets and to document the main data sources. In the Estimates Summary table, we enter the estimate and uncertainty range for each estimate. Table 1 shows whether to enter a number or percentage for each step.

Table 1.
Entering data in the Estimates Summary

Estimate	How to enter it
Step 2: initial population size estimate (PSE)	Enter initial size estimate as a positive number greater than 0 (e.g. 10 807)
Step 3: PSE projected to current year	Do not enter data directly in these cells The estimate and range will update automatically by pulling in the projected PSE calculated on the Project PSE tab
Step 4: proportion HIV-negative	Enter a percentage between 0% and 100%
Step 5: subgroup inflation factor	Enter a percentage greater than 0% For example, an inflation factor of 300% will triple the size estimate; an inflation factor of 50% will reduce the size estimate by half If no inflation is desired, enter 100% for the estimate, lower limit and upper limit Do not enter 0% or leave these cells blank
Step 6: proportion at risk	Enter a percentage between 0% and 100%

Be sure to document the data source in the columns to the right of each step.

Figures 1 and 2 show where to enter the risk proportion, depending on whether you are using a single proportion or low, medium and high risk proportions.

If using a single risk proportion, enter it in Step 6 (Figure 1).

Figure 1.

Example of Estimates Summary when using a single risk proportion

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES			
<i>Enter the point estimate and limits for each step. On the right, describe the scope of the data age range. Then compare the scope of the data sources used with the intended scope of the</i>			
	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	4,999	4,634	5,366
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	6,516	6,040	6,995
STEP 4. PROPORTION HIV-NEGATIVE	91.8%	86.8%	96.8%
STEP 5. SUBGROUP INFLATION FACTOR	142.9%	127.1%	163.1%
STEP 6. PROPORTION AT RISK	14.6%	4.1%	25.0%
ESTIMATED NUMBER AT RISK	1,248	334	2,162

Enter your risk proportion here

Don't edit the targets. They calculate automatically.

If using low, medium and high risk proportions, leave Step 6 blank (Figure 2), and enter the proportions in the next table down (Estimates by Risk Category). In either case, the targets and their ranges update automatically.

Figure 2.

Example of Estimates Summary when using low, medium and high risk proportions

	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	4,999	4,634	5,366
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	6,516	6,040	6,995
STEP 4. PROPORTION HIV-NEGATIVE	91.8%	86.8%	96.8%
STEP 5. SUBGROUP INFLATION FACTOR	142.9%	127.1%	163.1%
STEP 6. PROPORTION AT RISK	14.6%	4.1%	25.0%
ESTIMATED NUMBER AT RISK			

ESTIMATES BY RISK CATEGORY			
<i>Fill in the box below only if using the Risk Factors Approach to define risk with more than one risk category.</i>			
	Estimate	Lower limit	Upper limit
Risk proportions:			
Risk Category 1	14.6%	4.1%	25.0%
Risk Category 2	25.0%	15.7%	34.4%
Risk Category 3	60.4%	50.8%	70.0%
Estimated numbers at risk:			
Risk Category 1	1,448	334	2,100
Risk Category 2	2,136	1,271	3,002
Risk Category 3	5,162	4,013	6,310

Enter your risk proportions here

Don't edit the targets

The examples that follow illustrate the remaining worksheets:

- ▶ Project PSE sheet.
- ▶ Risk Proportion sheet.
- ▶ Minimum Behaviors Calculator when using the exposures approach to define risk criteria.

Example 1: Men who have sex with men—exposures approach

This example estimates the number of men who have sex with men at risk in City A in eastern Africa as of 2018. Survey data are from a respondent-driven sampling (RDS) biobehavioural surveillance (BBS) that included men who have sex with men and transgender women participants.

Step 1: Define the population

At the top of the PrEP Estimates sheet we enter an operational definition of the population. The aim is to estimate the number at risk among all males who currently engage in sex with other males. In this example, we restrict our definition to males aged 18 years and over who have had sex with another male in the past 6 months, definitions used in recent surveys and PSEs.

STEP 1. DEFINE THE GEOGRAPHIC AREA AND POPULATION FOR TARGET-SETTING

Indicate the intended scope of the estimates:

Population:	MSM
Operational definition of population:	Males who have had sex with another male in the past 6 months
Age range	18 and older
Geographic area:	City A
Desired year of the estimate	2018

As we complete the remaining steps, we will be alert to possible misalignment of the data sources.

Step 2: Initial PSE

Moving down the Estimates Summary table, we enter details of the initial PSE. In City A, there was 1 recent high-quality PSE conducted using 3-source capture–recapture among men who have sex with men at venues in 2017, with an estimate of 10 807 (95% CI 7371–14 244) men who have sex with men.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data sources used in terms of geography, time period, definition of the population, and age range. Then compare the scope of the data sources used with the intended scope of the estimates (defined in Step 1) in order to identify any alignment issues.

	ENTER ESTIMATES			DOCUMENT THE DATA SOURCE FOR EACH ESTIMATE. IDENTIFY ANY ALIGNMENT ISSUES.				
				Compare representativeness of each data source with the intended scope of the estimates (from Step 1)				
	Estimate	Lower limit	Upper limit	What is the source of this estimate?	Geographic area	Year(s) when data were collected	How the population was defined	Age range
STEP 2. INITIAL PSE	10 807	7 371	14 244	Three-source, venue-based capture-recapture using unique objects	City A	2017	Had anal sex with a male; encountered at an MSM venue	≥ 18 years

Step 3: Project the PSE

Because a year has passed since the initial PSE was conducted, we project the PSE forward to the desired year of implementation, 2018, using a projection calculator (Project PSE sheet). Since there is no census denominator available for males aged 15–49 years in City A in 2018, we choose projection method C to apply an annual growth rate to the initial PSE.

METHODS

USE THIS METHOD IF:

A

The percent of the KP or priority population estimated by the initial PSE is available. A census denominator for the desired year of implementation ("current year") is available.

B

Only the number of individuals estimated by the initial PSE is available and...

Census denominators are available for the year when the PSE was conducted and for the current year. Typically, the denominator is either the male or female population aged 15 to 49 years.

C

Annual growth rates are available, starting from the year when the PSE was conducted to the year just before the current year (e.g., last year).

Enter the method to be used:

C

The annual urban growth rate for the country in 2017 was 5.4% (or 1.054 as a proportion), obtained from the World Bank.

METHOD C

Apply annual growth rates to the number of population members estimated by the initial PSE.	Estimate	Lower limit	Upper limit
Initial PSE as an absolute number of individuals	10 807	7 371	14 244
Annual growth rates between year the PSE was conducted and the current year (expressed as "1 + the proportion", e.g., 1.15)	<i>Enter growth rates as needed. Leave all unused cells as 1.0.</i>		
Annual growth rate - Year 1	1.054	1.054	1.054
Annual growth rate - Year 2	1.000	1.000	1.000
Annual growth rate - Year 3	1.000	1.000	1.000
Annual growth rate - Year 4	1.000	1.000	1.000
Annual growth rate - Year 5	1.000	1.000	1.000
Annual growth rate - Year 6	1.000	1.000	1.000
Annual growth rate - Year 7	1.000	1.000	1.000
Annual growth rate - Year 8	1.000	1.000	1.000
Annual growth rate - Year 9	1.000	1.000	1.000
Projected PSE	11 391	7 769	15 013

The projection of 11 391 (95% CI 7769–15 013) men who have sex with men is copied automatically to the Estimates Summary table. We document the details of the data source on the same row.

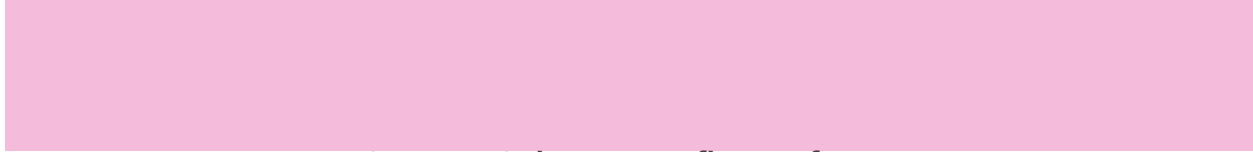
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	11 391	7 769	15 013	World Bank annual urban population growth rate	All urban areas of the country	2016 annual growth rate (assumed same for 2017)	General population	All ages
---	--------	-------	--------	--	--------------------------------	---	--------------------	----------

Step 4: Proportion HIV-negative

We estimate the proportion of our target population, men who have sex with men, that is HIV-negative from an RDS survey. The survey was among both men who have sex with men and transgender women, so we estimate the proportion among the subsample of participants who did not identify as transgender women.



STEP 4. PROPORTION HIV-NEGATIVE	88.6%	82.8%	94.3%	RDS IBBS survey of MSM & TW, subset to MSM participants	City A	2012-2013	Male sex; had sex with a male partner in the past 6 months	≥ 18 years
--	-------	-------	-------	---	--------	-----------	--	------------

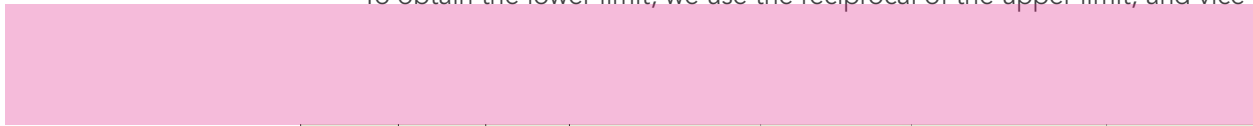


Step 5: Subgroup inflation factor

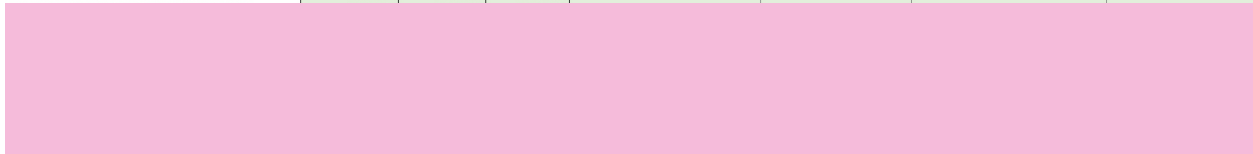
Because the PSE was venue-based and we want estimates for the larger population of men who have sex with men, we complete Step 5 to inflate the PSE to non-venue-going men who have sex with men. Fortunately, the same RDS survey included a question about frequenting venues, which can serve as the inflator.

Among participants who were HIV-negative, the proportion who responded that they frequented the venues included in the PSE (clubs, bars, hotels, cafes, restaurants) was 40.2%. We enter 1 divided by this estimate (the reciprocal), or 248.8%, as the inflation factor.

To obtain the lower limit, we use the reciprocal of the upper limit, and vice versa.



STEP 5. SUBGROUP INFLATION FACTOR	248.8%	194.2%	347.2%	RDS IBBS survey of MSM & TW, subset to HIV-, MSM participants: Question item, "In general, where do you usually meet men you have sex with?" (Responses club, bar, hotel, café or restaurant)	City A	2012-2013	Male sex; had sex with a male partner in the past 6 months	≥ 18 years
--	--------	--------	--------	---	--------	-----------	--	------------



Step 6: Risk proportion

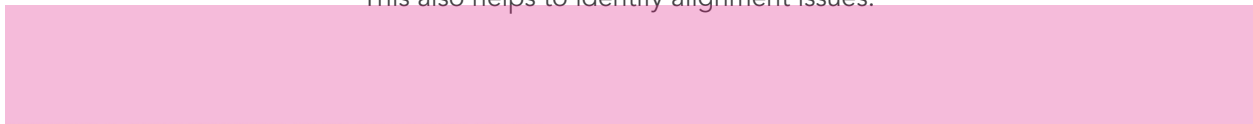
Step 6A: Choose representative data source

We select the RDS survey of men who have sex with men and transgender women as the data source, since it is the most representative and recent BBS available in City A.

Step 6B: Define substantial risk of HIV infection

We use the exposures approach to define the risk criteria and estimate the proportion of men who have sex with men at risk. In this case, we can use the exposures approach because the RDS survey included questions on the number of male partners and the number of unprotected anal sex acts with male partners recently (past six months).

We fill out Section A of the Risk Proportion sheet to document the survey population. This also helps to identify alignment issues.



SECTION A. WHAT SURVEY DATA ARE BEING USED?

Fill in details about the survey that will be used to estimate risk behaviors.

Population surveyed:	MSM
Operational definition in survey:	Males who have had sex with another male in the past 6 months
Age range or eligibility criteria	18 and older
Geographic area:	City A
Sampling method:	RDS
Year of survey:	2012-2013



Step 6Bi: Set target incidence

We use the WHO-recommended target incidence of 3% since there has been no local study recommending a different incidence threshold for PrEP. We enter our target into the Minimum Behaviors Calculator.

MINIMUM BEHAVIORS CALCULATOR

Risk Proportion Tool

Step 6. Defining Risk, Exposure Approach

MSM

INSTRUCTIONS:

This worksheet produces a table of minimum levels of risk be participants as "at risk" of HIV. *Enter in the incidence target*

LEGEND

Enter data in the green cells

Blue cells calculate automatically

PARAMETERS

Percent

HIV incidence target

3%

Step 6Bii: Decide which profiles to use

Back on the Risk Proportions A sheet, we find out which profiles we can use by filling in the first two items under Section B (sexually transmitted infection status of participants and whether participants know about any partners' HIV status).

We do not necessarily need these data to use the exposures approach, but we are fortunate to have roughly these data in the survey, so we enter "yes". The table updates to show us we can use profiles 1, 2 and 3.

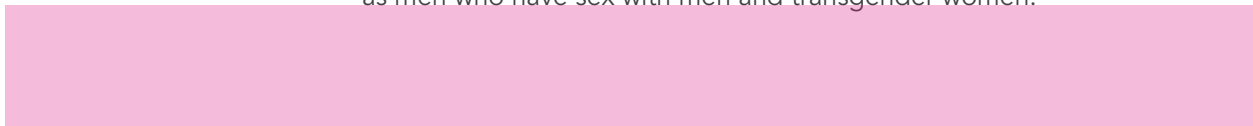
Because only one sexually transmitted infection (syphilis) was assessed by biological testing in the survey, we choose to define sexually transmitted infection broadly as either a positive syphilis result or a self-report of recent sexually transmitted infection symptoms.

Data item	Is this data item available? (Yes or No)	How was this item defined? <i>(fill in only if the item is available)</i>
1. STI status of survey participants	yes	Test positive for syphilis; or past 6-months STI sign or symptom
2. Whether participant has any recent known HIV+ partners	yes	Stable or casual partner in past 6 months has HIV
Profiles that can be used with the data available:	1, 2, & 3	

Step 6Biii: Determine partner population characteristics

To use the Risk Proportions A sheet, we need prevalence estimates of HIV and sexually transmitted infection in the partner population.

Partners of men who have sex with men could be defined either as men who have sex with men or as men who have sex with men transgender women. Because our RDS survey included men who have sex with men and transgender women, it is easy for us to develop estimates for this combined population. We define the partner population as men who have sex with men and transgender women.



<p>Partner characteristics: <i>partners are MSM or MSM+TW</i></p>		
<p>5. HIV prevalence in the partner population</p>		<p>Prevalence of HIV from RDS survey, subset to MSM participants, and including HIV+ and HIV- participants.</p>
<p>6. STI prevalence in the partner population</p>		<p>Prevalence of (syphilis or past 12-month STI symptoms) from RDS survey, subset to MSM participants, and including HIV+ and HIV- participants.</p>
<p>7. ART coverage of PLWH in the partner population</p>		<p>Assumed (but should be drawn from program data)</p>

We use RDS Analyst software with the RDS-2 estimator¹ on the full sample of men who have sex with men and transgender women, including both HIV-positive and HIV-negative participants. We estimate the prevalence of the derived sexually transmitted infection variable (a positive syphilis result or self-report of recent sexually transmitted infection symptoms). We estimate prevalence of HIV at 13.2% and of sexually transmitted infections at 54.9%

We assume that antiretroviral therapy coverage among men who have sex with men and sexually transmitted infection with HIV is 50%. This includes coverage of both diagnosed and undiagnosed HIV.

PARAMETERS	Percent
HIV incidence target	3%
<i>PARTNER POPULATION (MSM; or MSM & TW)</i>	
HIV prevalence	13%
STI prevalence	55%
ART coverage among PLWH	50%

Step 6Biv: Look up minimum number of partners and risky acts

After entering the target incidence and partner estimates into the Minimum Behaviors Calculator sheet, the table of minimum risk behaviours updates automatically.

The table shows that men who have sex with men in City A with a male partner living with HIV (Profile 1) who engage in at least 6 acts of unprotected anal intercourse per year would meet the 3% incidence target.

For men who have sex with men in Profiles 2 and 3, the minimum number of unprotected anal intercourse acts depends on the number of unprotected anal intercourse partners.

For participants in Profile 2 (with a sexually transmitted infection and without any reported partner living with HIV), those who had 1 unprotected anal intercourse partner per year would need at least 11 unprotected anal intercourse acts per year to reach 3% incidence. Those with two unprotected anal intercourse partners per year would need at least five unprotected anal intercourse acts per partner per year, and so on.

These results are based on a force of infection model embedded in the spreadsheet.

¹ We chose RDS-2 for this example, but other estimators are possible.

TABLE OF MINIMUM RISK BEHAVIORS TO MEET TARGET INCIDENCE

Table stops when minimum risky acts per partner is 1.

UAI partners per year	Minimum UAI acts per partner per year		
	Profile 1	Profile 2	Profile 3
1	6	11	19
2	-	5	8
3	-	4	5
4	-	3	4
5	-	2	3
6	-	2	3
7	-	2	2
8	-	2	2
9	-	1	2
10	-	-	2
11	-	-	2
12	-	-	2
13	-	-	2
14	-	-	1
15	-	-	-

Step 6C: Classify risk level of survey participants

The purpose of these minimum numbers is to determine whether each participant in the survey faces an HIV incidence of at least 3%. We need to compare reported numbers of acts and partners from the survey against these minimum thresholds. To do this, we create new variables in the dataset:

- ▶ Unprotected anal intercourse acts: the survey asked about the number of anal sex acts, and the number of anal sex acts with a condom, in the past six months. We create a new variable in the data that subtracts these two responses to obtain the number of unprotected anal intercourse acts.
- ▶ Unprotected anal intercourse partners: the survey did not ask specifically about number of recent unprotected anal intercourse partners. Instead, we use the total number of male sex partners, which we had to sum across reported numbers of steady, casual, paid and paying male sex partners in the past six months. We now have an inconsistency, because some participants reported more partners than unprotected anal intercourse acts because they may not have engaged in unprotected anal intercourse with all partners reported. As a correction, we cap the number of partners at the number of acts. This is not a perfect solution, but it may be the best we can do with these data.

- ▶ Unprotected anal intercourse acts per partner: we divide the number of unprotected anal intercourse acts by the number of unprotected anal intercourse partners to get the number of unprotected anal intercourse acts per partner.
- ▶ Checking skip patterns: we look for skip patterns in the questionnaire that could cause problems for these measures. We replace missing values of nacts and npartners with 0 in some cases. For example, some participants reported no anal sex or no unprotected anal intercourse in the past six months and so were not asked about the precise number. Thus, they need to be replaced by 0 so they can be correctly defined as not at risk in the risk proportion.
- ▶ Annualizing: because the data in the survey refer to the past 6 months and the Risk Proportions A sheet shows the minimum thresholds for the past 12 months, we annualize the responses in the survey. We multiply the number of risky partners and the number of risky acts per partner by $\sqrt{2}$ (= 1.41). (It may seem intuitive to multiply by 2, but this is incorrect.²)

We document how the measures were defined in Section B of the Risk Proportion sheet.

Participant risk behaviors:		
		No. steady, casual, paid, and paying male sex partners in past 6m. (There was no way to limit to this to UAI partners specifically).
3. No. UAI partners per year		#partners restricted to be no greater than #acts. Annualized by multiplying by sqrt(2)
		[(#rounds anal sex with a male in past 6months) - (#rounds anal sex in past 6 months in which a condom was used)] / (no. male sex partners in past 6 months).
4. No. UAI acts per partner per year		Annualized by multiplying by sqrt(2)

² If we try to annualize from the past 6 months to the past 12 months by doubling the number of partners, X, and the number of acts per partner, Y, this implies 4 times the total number of acts ($2X \times 2Y = 4XY$), which is not correct. Instead, to scale the reported numbers in a way that preserves the total number of acts over time, we multiply by $\sqrt{2}$. Then, the total acts works out correctly ($\sqrt{2}X \times \sqrt{2}Y = 2XY$). This is just one of many ways to annualize.

In Section C, we document how we used the minimum numbers to classify each participant as at risk. We generate new variables, called *nacts* and *npartners*, with the definitions above. Then we generate a variable, *atrisk*, which is coded as 1 if the acts and partners are at least as high as the thresholds and 0 if they are not. We set *atrisk* to missing if either of the *nacts* or *npartners* variables is missing, so participants with incomplete data are excluded from the risk proportion.

	PROFILE 1	PROFILE 2	PROFILE 3
Profile definition	Survey participant has a known HIV partner.	Survey participant has STI. HIV status of partners is unknown or was not assessed.	Survey participant does not have STI or STI was not assessed. HIV status of partners is unknown or was not assessed.
What rules or criteria are used to classify participants into this profile in the dataset?	A stable or casual partner in past 6m had HIV.	Tested positive for syphilis at survey or reported STI symptoms in past 6mo. Did not report having a stable or casual partner in past 6m with HIV.	Tested negative for syphilis at survey and did not report STI symptoms in past 6mo. Did not report having a stable or casual partner in past 6m with HIV.
What rules or criteria were used to determine if participants in this profile were "at risk"?	Any UAI in period (calculated threshold was 6 acts. However, more conservative definition used as number of acts specifically with HIV partner is unknown).	As in min. behaviors table, i.e.: 1 partner & ≥ 11 UAI acts 2 partners & ≥ 5 UAI acts 3 partners & ≥ 4 UAI acts 4 partners & ≥ 3 UAI acts 5-8 partners & ≥ 2 UAI acts ≥ 9 partners & ≥ 1 UAI act	As in min behaviors table, i.e.,-: 1 partner & ≥ 19 UAI acts 2 partners & ≥ 8 UAI acts 3 partners & ≥ 5 UAI acts 4 partners & ≥ 4 UAI acts 5-6 partners & ≥ 3 UAI acts 7-13 partners & ≥ 2 UAI acts ≥ 14 partners & ≥ 1 UAI act

Step 6D: Estimate proportion at risk among HIV-negative subjects

With the variable *atrisk* added to the dataset, we estimate the population proportion of *atrisk* using RDS Analyst, subset to participants who tested negative for HIV. We enter the estimate of 10.5% (95% CI 4.8–16.1%) into the Risk Proportions A sheet and then on the Estimates Summary of the PrEP Estimates sheet.

The estimated number at risk is calculated automatically on the Estimates Summary sheet as 2636 (95% CI 792–4480).

	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	10 807	7 371	14 244
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	11 391	7 769	15 013
STEP 4. PROPORTION HIV-NEGATIVE	88.6%	82.8%	94.3%
STEP 5. SUBGROUP INFLATION FACTOR	248.8%	194.2%	347.2%
STEP 6. PROPORTION AT RISK	10.5%	4.8%	16.1%
ESTIMATED NUMBER AT RISK	2 636	792	4 480

Example 2: Men who have sex with men—risk factors approach

This example takes up the scenario from Example 1 but uses the risk factors approach instead of the exposures approach to define risk.

Steps 1–5 are the same as in Example 1. Here we illustrate only Step 6.

Step 6: Risk proportion

Step 6A: Choose a representative data source

As in Example 1, we draw on an RDS survey of men who have sex with men and transgender women.

Step 6B: Define substantial risk of HIV infection

We use the risk factors approach to define the risk criteria. We draw on three established risk factors associated with increased risk of HIV infection among men who have sex with men and are available in our survey: sexually transmitted infection; receptive role in anal sex; and unprotected anal intercourse.

We defined three levels of risk (or risk categories) based on these risk factors:

- ▶ High-risk: positive syphilis result or self-report of sexually transmitted infection symptoms in past 12 months.
- ▶ Medium-risk: receptive and any unprotected sex in past 12 months.
- ▶ Low-risk: all other participants who do not meet either of the above criteria.

These definitions are not derived from any model as in the exposures approach. They are subjective but based on risk factors identified in the literature. Many other potential categorization schemes are possible.

As in the exposures approach, we fill out Section A of the Risk Proportion sheet to document the population included in the survey.

SECTION A. WHAT SURVEY DATA ARE BEING USED?

Fill in details about the survey that will be used to estimate risk behaviors.

Population surveyed:	MSM
Operational definition in survey:	Males who have had sex with another male in the past 12 months
Age range or eligibility criteria	18 and older
Geographic area:	Guatemala City, Guatemala
Sampling method:	RDS
Year of survey:	2016

Step 6C: Classify risk level of survey participants

We use responses to question items to create measures of sexually transmitted infection, receptive anal sex, and any unprotected anal intercourse in the past 12 months in our data.

Before we classify participants into different risk levels, we look for skip patterns in the questionnaire that could cause problems in the three risk factor measures. This leads us to replace missing values with 0 in some cases. For example, some participants reported no anal sex at all in the past 12 months and so were not asked questions about receptive role in anal sex at all; we assigned them a value of 0 for the receptive anal sex measure. There were similar issues with the other two measures.

We create a categorical variable, *risk_level*, coded 1, 2 or 3 for the respective risk categories, being careful with missing values. We code *risk_level* as missing if the risk level is unclear due to non-response, so participants with incomplete data are excluded from the estimates.

In Sections D and E of the Risk Proportion sheet, we document how the risk factors and risk categories were defined.

SECTION D. HOW WERE THE RISK CATEGORIES DEFINED? WHAT PART OF THE POPULATION IS IN EACH CATEGORY?

Describe the risk criteria and the measures used to define risk. Three categories are listed. However, you may use one or two categories or as many as needed.

	RISK CATEGORY 1	RISK CATEGORY 2	RISK CATEGORY 3
How was this risk category defined?	Syphilis or STI symptoms in the past 12 months	Not in risk category 1 Receptive and any UAI in past 12 months	All others
What percent of all MSM are in this risk category?	14.60%	25.00%	60.40%
Lower limit	4.10%	15.70%	50.80%
Upper limit	25.00%	34.40%	70.00%

Copy the risk proportions above onto the PrEP Estimates worksheet (Step 6).

SECTION E. WHAT MEASURES WERE USED TO DEFINE THE RISK CATEGORIES?

Describe any measures utilized to define the risk categories above.

	Name of measure	How was this measure defined?
Measure 1	STI	Test positive for syphilis; or past 12-month STI symptoms
Measure 2	Receptive	Engaged primarily in receptive or receptive & insertive anal sex with any type of partner (stable, casual, or commercial) over the past 12 months
Measure 3	Any UAI	"Sometimes" or "never" (vs. "always") used a condom with any partner type over the past 12 months
Measure 4		
Measure 5		

Step 6D: Estimate proportion at risk among HIV-negative subjects

With the categorical variable *risk_level* added to the dataset, we estimate the population proportions of each level using RDS Analyst, subset to participants who tested negative for HIV. The point estimates were 14.6%, 25.0% and 60.4%. We enter these estimates in Section E along with their 95% CIs (shown above).

These estimates are copied automatically to the Estimates Summary table on the PrEP Estimates sheet. The estimated number at risk in each risk category updates automatically.

Because we have more than one risk category, we leave the Step 6 row blank on the Estimates Summary table and enter the proportions into the Estimates by Risk Category table just below it.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data source; compare the scope of the data sources used with the intended scope of the estimates (defined in St

	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	4 999	4 634	5 366
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	6 516	6 040	6 995
STEP 4. PROPORTION HIV-NEGATIVE	91.8%	86.8%	96.8%
STEP 5. SUBGROUP INFLATION FACTOR	142.9%	127.1%	163.1%
STEP 6. PROPORTION AT RISK			
ESTIMATED NUMBER AT RISK	0	0	0

ESTIMATES BY RISK CATEGORY (APPROACH III ONLY)

Fill in the box below only if using Approach III ("Markers") to define risk with more than one risk category.

	Estimate	Lower limit	Upper limit
Risk proportions:			
Risk Category 1	14.6%	4.1%	25.0%
Risk Category 2	25.0%	15.7%	34.4%
Risk Category 3	60.4%	50.8%	70.0%
Estimated numbers at risk:			
Risk Category 1	1 248	334	2 162
Risk Category 2	2 136	1 271	3 002
Risk Category 3	5 162	4 013	6 310

The estimates suggest that as of 2017 in Guatemala City, there were approximately 1248 (95% CI 334–2162) men who have sex with men in risk category 1; 2136 (1271–3002) in risk category 2; and 5162 (4013–6310) in risk category 3.

Example 3: Female sex workers—exposures approach

This example estimates the number of female sex workers at risk in City A in eastern Africa as of 2018. Survey data are from an RDS BBS.

Step 1: Define the population

At the top of the PrEP Estimates sheet, we enter an operational definition of the population. Our aim is to estimate the number at risk among all females who currently sell sex to male clients. We restrict our definition to females aged 15 years and over who sell sex, based on the definition used in the best PSE available.

STEP 1. DEFINE THE GEOGRAPHIC AREA AND POPULATION FOR TARGET-SETTING

Indicate the intended scope of the estimates:

Population:	FSW
Operational definition of population:	Males who have had sex with another male in the past 6 months
Age range	15 and older
Geographic area:	City A
Desired year of the estimate	2018

As we complete the remaining steps, we will be alert to possible misalignment of the data sources.

Step 2: Initial PSE

Moving down the Estimates Summary table, we enter details of the initial PSE. For female sex workers in City A, there was 1 recent high-quality PSE conducted using 3-source capture–recapture among female sex workers at identifiable sex work locations in 2017, with an estimate of 15 419 (95% CI 10 818–20 021) female sex workers.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data sources used in terms of geography, time period, definition of the population, and age range. Then compare the scope of the data sources used with the intended scope of the estimates (defined in Step 1) in order to identify any alignment issues.

	ENTER ESTIMATES			DOCUMENT THE DATA SOURCE FOR EACH ESTIMATE. IDENTIFY ANY ALIGNMENT ISSUES.				
	Estimate	Lower limit	Upper limit	Compare representativeness of each data source with the intended scope of the estimates (from Step 1)				
				What is the source of this estimate?	Geographic area	Year(s) when data were collected	How the population was defined	Age range
STEP 2. INITIAL PSE	15 419	10 818	20 021	Three-source, venue-based capture-recapture using unique objects	City A	2017	Female actively selling sex to men or women (in places where FSW wait for clients)	≥ 15 years



Step 3: Project the PSE

Because a year has passed since the initial PSE was conducted, we project the PSE forward to the desired year of implementation, 2018, using a projection calculator (Project PSE sheet). Since there is no census denominator available for females aged 15–49 years in City A in 2018, we choose projection method C to apply an annual growth rate to the initial PSE.



METHODS

USE THIS METHOD IF:

A	The <u>percent</u> of the KP or priority population estimated by the initial PSE is available. A census denominator for the desired year of implementation ("current year") is available.
B	Only the <u>number</u> of individuals estimated by the initial PSE is available and... Census denominators are available for the year when the PSE was conducted and for the current year. Typically, the denominator is either the male or female population aged 15 to 49 years.
C	Annual growth rates are available, starting from the year when the PSE was conducted to the year just before the current year (e.g., last year).

Enter the method to be used:

C



The annual urban growth rate for African Country A in 2017 was 5.4% (or 1.054 as a proportion), obtained from the World Bank.

METHOD C

Apply annual growth rates to the number of population members estimated by the initial PSE.	Estimate	Lower limit	Upper limit
Initial PSE as an absolute number of individuals	15 419	10 818	20 021
Annual growth rates between year the PSE was conducted and the current year (expressed as "1 + the proportion", e.g., 1.15)	<i>Enter growth rates as needed. Leave all unused cells as 1.0.</i>		
Annual growth rate - Year 1	1.054	1.054	1.054
Annual growth rate - Year 2	1.000	1.000	1.000
Annual growth rate - Year 3	1.000	1.000	1.000
Annual growth rate - Year 4	1.000	1.000	1.000
Annual growth rate - Year 5	1.000	1.000	1.000
Annual growth rate - Year 6	1.000	1.000	1.000
Annual growth rate - Year 7	1.000	1.000	1.000
Annual growth rate - Year 8	1.000	1.000	1.000
Annual growth rate - Year 9	1.000	1.000	1.000
Projected PSE	16 252	11 402	21 102

The projection, 16 252 (95% CI 11 402–21 102) female sex workers, is copied automatically to the Estimates Summary table. We document the details of the data source on the same row.

STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	16 252	11 402	21 102	World Bank annual urban population growth rate	All urban areas of African country A	2016 annual growth rate (assumed same for 2017)	General population	All ages
---	--------	--------	--------	--	--------------------------------------	---	--------------------	----------

Step 4: Proportion HIV-negative

We estimate the proportion of our target population, female sex workers, that is HIV-negative from an RDS survey.

				RDS IBBS survey of FSW	City A	2012-2013	Female. Sold sex in the past 6 months for money, goods or services.	≥ 15 years
STEP 4. PROPORTION HIV-NEGATIVE	68.2%	64.6%	71.8%					

Step 5: Subgroup inflation factor

Because the PSE was venue-based and we want estimates for the larger population of female sex workers, we complete Step 5 to inflate the PSE to female sex workers who may not be found at identifiable sex work sites.

Fortunately, the same RDS survey included a question about how female sex workers meet their clients, which can serve as the inflator.

Among female sex workers participants who were HIV-negative, the proportion who responded that they sold sex at the kinds of venue included in the PSE (streets, hotels, lodges, clubs, bars, restaurants, brothels) was 78.4%. In the sheet, we enter 1 divided by this estimate (the reciprocal), or 127.6%, as the inflation factor. To obtain the lower limit, we use the reciprocal of the upper limit, and vice versa.

				RDS IBBS survey of FSW, subset to HIV- participants: Question item, "In general, where do you usually meet clients you sell sex to?" (Responses street, hotel, lodge, club, bar, restaurant, or brothel; this inflates to include phone, sms, email, internet, and 'private place')	City A	2012-2013	Female. Sold sex in the past 6 months for money, goods or services.	≥ 15 years
STEP 5. SUBGROUP INFLATION FACTOR	127.6%	120.1%	136.1%					

Step 6: Risk proportion

Step 6A: Choose a representative data source

We select the RDS survey of female sex workers as the data source, since it is the most representative and recent BBS available in City A.

Step 6B: Define substantial risk of HIV infection

We use the exposures approach to define the risk criteria and estimate the proportion of female sex workers at risk. In this case, we can use the exposures approach because the RDS survey included questions on the number of male clients, and the number of sex acts with clients, in which a condom was used recently (past six months). We use these data to calculate the number of sex acts in which a condom was not used (number of risky acts).

We fill out Section A of the Risk Proportion sheet to document the survey population. This also helps to identify alignment issues.

STEP 1. DEFINE THE GEOGRAPHIC AREA AND POPULATION FOR TARGET-SETTING

Indicate the intended scope of the estimates:

Population:	FSW
Operational definition of population:	Males who have had sex with another male in the past 6 months
Age range:	15 and older
Geographic area:	City A
Desired year of the estimate:	2018

MINIMUM BEHAVIORS CALCULATOR

Risk Proportion Tool

Step 6. Defining Risk, Exposure Approach

FSW

INSTRUCTIONS:

This worksheet produces a table of minimum levels of risk by participants as "at risk" of HIV. Enter in the incidence target

LEGEND

Enter data in the green cells

Blue cells calculate automatically

PARAMETERS

HIV incidence target

Percent

2.0%

Step 6Bi: Set target incidence

For this example, we use a target incidence of 2% to demonstrate it is possible to use any value. In practice, 3% (the WHO-recommended value) should be used unless a local study recommends a different threshold. We enter our target into the Minimum Behaviors Calculator.

Step 6Bii: Decide which profiles to use

On the Risk Proportions Summary sheet, we find out which profiles we can use by filling in the first two items under Section B (sexually transmitted infection status of participants and whether participants know about any partners' HIV status). Our RDS survey did not collect information about whether clients had HIV (that is difficult to know), but we can measure sexually transmitted infection as a positive syphilis test result or self-reported sexually transmitted infection symptoms. We enter "yes" under sexually transmitted infection (Item 1) and "no" under HIV status of partners (Item 2). The table updates to show us we can use Profiles 2 and 3.

Data item	Is this data item available? (Yes or No)	How was this item defined? <i>(fill in only if the item is available)</i>
1. STI status of survey participants	yes	Test positive for syphilis; or past 6-month STI sign or symptom (lesions, sores, or warts on private parts, or pain when urinating)
2. Whether participant has any recent known HIV+ partners	no	
Profiles that can be used with the data available:	2 & 3	

Step 6Biii: Determine partner population characteristics

To use the Minimum Behaviors Calculator, we need prevalence estimates of HIV and sexually transmitted infections in the partner population. We define partners as male sex clients only.

Partner characteristics: <i>partners are male sex clients</i>		
5. HIV prevalence in the partner population		Prevalence of HIV among males in urban areas of African country A, from UPHIA 2016-2017
6. STI prevalence in the partner population		Prevalence of active syphilis among adults ages 15-64 in urban areas of African country A, from UPHIA 2016-2017
7. ART coverage of PLWH in the partner population		Percent of all PLWH who have suppressed viral loads, from UNAIDS African country A factsheet 2017

We do not have data on HIV prevalence, sexually transmitted infection prevalence or antiretroviral therapy coverage for male sex clients specifically, or for males in City A.

Instead, we use data on the general male population nationally from the African Country A population-based HIV impact assessment (PHIA) in 2016–2017. PHIA estimated 4.6% HIV prevalence among males in urban areas of African Country A. For sexually transmitted infection, we use the Uganda PHIA estimate of 1.6% active syphilis among all adults aged 15–64 years in urban areas. This estimate includes females, but the report notes that rates were similar among males and females.

For antiretroviral therapy coverage, we draw on the UNAIDS country factsheet for African Country A for 2017, which estimates 56% of all people living with HIV have suppressed viral loads. This estimate could be improved by using an estimate specific to males or, even better, to males in City A.

PARTNER POPULATION (male sex clients)

HIV prevalence	4.6%
STI prevalence	1.6%
ART coverage among PLWH	56.0%

Step 6Biv: Look up minimum number of partners and risky acts

After entering the target incidence and partner estimates into the Minimum Behaviors Calculator, the table of minimum risk behaviours updates automatically. We ignore the Profile 1 column because we are using only Profiles 2 and 3.

The table shows the number of risky partners in the first column (male sex clients with whom female sex workers engaged in condomless sex in the past year).

To meet the 2% incidence target, the table shows that female sex workers in Profile 2 (with sexually transmitted infection) who had 1 risky partner per year would need to have condomless sex with that partner at least 447 times during the year. Female sex workers with 2 risky partners per year would need to have condomless sex at least 131 times with each risky partner per year, and so on.

Female sex workers in Profile 3 (no sexually transmitted infection) who had 1 or 2 risky partners per year would need to have more than 500 condomless sex acts per partner to meet the 2% incidence threshold.

The table does not list specific values greater than 500. Because 500 acts would imply more than one act per day per partner, it is assumed that individuals do not reach this limit.

Female sex workers in Profile 3 with 3 risky partners per year would need at least 363 condomless sex acts per partner per year. Those with 4 risky partners per year would need at least 258 condomless sex acts per partner per year, and so on.

These results are based on a force of infection model embedded in the spreadsheet.

TABLE OF MINIMUM RISK BEHAVIORS TO MEET TARGET INCIDENCE

Table stops when minimum risky acts per partner is 1.

Condomless sex partners per year	Minimum condomless sex acts per partner per year		
	Profile 1	Profile 2	Profile 3
1	45	447	>500
2	-	131	>500
3	-	78	363
4	-	56	258
5	-	44	201
6	-	36	164
7	-	31	139
8	-	27	120
9	-	24	106
10	-	21	95
11	-	19	86
12	-	18	79
13	-	16	72
14	-	15	67
15	-	14	62
16	-	13	58
17	-	12	55
18	-	12	52
19	-	11	49
20	-	11	46
21	-	10	44
22	-	10	42
23	-	9	40
24	-	9	39
25	-	9	37
26	-	8	36
27	-	8	34
28	-	8	33
29	-	7	32
30	-	7	31

Step 6C: Classify risk level of survey participants

The purpose of these minimum numbers is to determine whether each participant in the survey faces an HIV incidence of at least 2%. We need to compare reported numbers of acts and partners from the survey against these minimum thresholds. To do this, we create new variables in the dataset:

- ▶ Risky acts: the survey asked about the number of sex acts with male clients, and the number of those acts in which a condom was used, in the past six months. We create a new variable in the data that subtracts these two responses to obtain the number of condomless sex acts.
- ▶ Risky partners: the survey asked about the total number of male sex clients, but it did not ask specifically about the number of risky clients (condomless sex). We calculate this by assuming that the percentage of clients who were risky was the same as the percentage of sex acts that were risky:
 - ▶ We calculate the percentage of sex acts where a condom was used, p .
 - ▶ To get the number of clients that were risky, we multiply the total number of clients by p .

This is an assumption that should be discussed and agreed upon with local experts. To avoid inconsistencies between the number of acts and the number of partners, we cap the number of partners at the number of acts. (A different approach would be to exclude any inconsistent records from the calculations.)

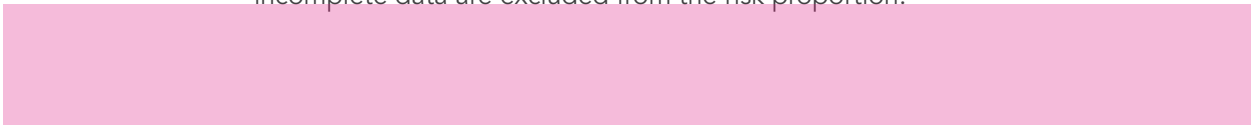
- ▶ Risky acts per partner: we divide the number of risky acts by the number of risky partners to get the number of risky acts per partner.
- ▶ Checking skip patterns: we look for skip patterns in the questionnaire that could cause problems for these measures. This leads us to replace missing values with 0 in some cases. For example, some participants reported no condomless sex in the past six months and so were not asked about the precise number. We replace these with 0 so they can be correctly defined as not at risk in the risk proportion.
- ▶ Annualizing: because the data in the survey refer to the past 6 months and the Minimum Behaviors Calculator shows the minimum thresholds for the past 12 months, we annualize the responses in the survey. We do this by multiplying both the number of risky partners and the number of risky acts per partner by $\sqrt{2}$ (= 1.41) (as explained earlier).

We document how the measures were defined in Section B of the Risk Proportion sheet.

Participant risk behaviors:		
3. No. condomless sex partners per year		No. male sex clients in past 6m. (there was no way to limit this to condomless sex partners specifically), #partners restricted to be no greater than #acts. Annualized by multiplying by sqrt(2)
4. No. condomless sex acts per partner per year		[(#rounds sex with a male in past 6months) - (#rounds sex in past 6 months in which a condom was used)] / (no. male sex partners in past 6 months). Annualized by multiplying by sqrt(2)



In Section C, we document how we used the minimum numbers to classify each participant as at risk. We generate new variables, called *nacts* and *npartners*, with the definitions above. We then generate a variable, *atrisk*, coded as 1 if the acts and partners are at least as high as the thresholds and 0 if they are not. We set *atrisk* to missing if either of the *nacts* or *npartners* variables was missing, so participants with incomplete data are excluded from the risk proportion.



	PROFILE 1	PROFILE 2	PROFILE 3
Profile definition	Survey participant has a known HIV partner.	Survey participant has STI. HIV status of partners is unknown or was not assessed.	Survey participant does not have STI or STI was not assessed. HIV status of partners is unknown or was not assessed.
What rules or criteria are used to classify participants into this profile in the dataset?		Tested positive for syphilis at survey or reported STI symptoms in past 6mo.	Tested negative for syphilis at survey and did not report STI symptoms in past 6mo.
What rules or criteria were used to determine if participants in this profile were "at risk"?		As in min. behaviors table, i.e.: 1 male client & ≥ 447 condomless sex acts 2 male clients & ≥ 131 condomless sex acts 3 male clients & ≥ 78 condomless sex acts 4 male clients & ≥ 56 condomless sex acts 5 male clients & ≥ 44 condomless sex acts 6 male clients & ≥ 36 condomless sex acts 7 male clients & ≥ 31 condomless sex acts etc.	As in min behaviors table, i.e.: 3 male clients & ≥ 363 condomless sex acts 4 male clients & ≥ 258 condomless sex acts 5 male clients & ≥ 201 condomless sex acts 6 male clients & ≥ 164 condomless sex acts 7 male clients & ≥ 139 condomless sex acts etc.

Step 6D: Estimate proportion at risk among HIV-negative subjects

With the variable *atrisk* added to the dataset, we estimate the population proportion of *atrisk* using RDS Analyst, subset to female sex worker participants who tested negative for HIV. We enter the estimate of 0.91% (95% CI 0.0–2.1%) into the Risk Proportions A sheet and on the Estimates Summary of the PrEP Estimates sheet.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data source range. Then compare the scope of the data sources used with the intended scope of the estimate.

	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	15,419	10,818	20,021
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	16,252	11,402	21,102
STEP 4. PROPORTION HIV-NEGATIVE	68.2%	64.6%	71.8%
STEP 5. SUBGROUP INFLATION FACTOR	127.6%	120.1%	136.1%
STEP 6. PROPORTION AT RISK	0.9%	0.0%	2.1%
ESTIMATED NUMBER AT RISK	129	0	283

The estimated number at risk is calculated automatically on the Estimates Summary sheet as 129 (95% CI 0–283).

Example 4: Transgender women—risk factors approach

This example estimates the number of transgender women at risk in Guayaquil, Ecuador as of 2018. Survey data are from an RDS BBS survey of transgender women, but we analyse the data as a convenience sample because it did not meet the planned sample size. The BBS included a unique object multiplier PSE that we use for the estimates.

Step 1: Define the population

At the top of the PrEP Estimates sheet, we enter an operational definition of the population. Our aim is to estimate the number at risk among all transgender women who currently engage in sex with other males. We restrict our definition to people aged 18 years and over who were assigned male sex at birth, who identify with the female gender, and who have had anal sex in the past 12 months with a male partner, to match the definition used in the 2017 BBS.

STEP 1. DEFINE THE GEOGRAPHIC AREA AND POPULATION FOR TARGET-SETTING

Indicate the intended scope of the estimates:

Population:	Transwomen
Operational definition of population:	Persons assigned male sex at birth who describe their gender as female
Age range:	18 and older
Geographic area:	Guayaquil, Ecuador
Desired year of the estimate:	2018

As we complete the remaining steps, we will be alert to possible misalignment of the data sources.

Step 2: Initial PSE

Moving down the Estimates Summary table, we enter details of the initial PSE. There are two potential PSEs to choose from for transgender women in Guayaquil: a unique object multiplier PSE in 2017, and a venue-based programmatic mapping in 2015.

The two PSEs produced similar estimates, so we use the multiplier PSE without further consideration of which may be the stronger estimate. We enter the estimate of 2410 (95% CI 1720–3556) on the row for Step 2, along with details of the PSE methods.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data sources used in terms of geography, time period, definition of the population, and age range. Then compare the scope of the data sources used with the intended scope of the estimates (defined in Step 1) in order to identify any alignment issues.

	ENTER ESTIMATES			DOCUMENT THE DATA SOURCE FOR EACH ESTIMATE. IDENTIFY ANY ALIGNMENT ISSUES.				
	Estimate	Lower limit	Upper limit	Compare representativeness of each data source with the intended scope of the estimates (from Step 1)				
				What is the source of this estimate?	Geographic area	Year(s) when data were collected	How the population was defined	Age range
STEP 2. INITIAL PSE	2,410	1,720	3,556	Multiplier estimate based on an RDS survey and distribution of unique objects	Guayaquil	2017	Persons assigned male sex at birth who describe their gender as female	≥ 18 years

Step 3: Project the PSE

Because a year has passed since the initial PSE was conducted, we project the PSE forward to 2018 using the projection calculator (Project PSE sheet).

We use projection Method B. We enter census projections for the general male population for the year of the PSE (2017) and for the current year (2018). The spreadsheet draws on these projections to calculate the growth rate in the general population during this period and uses it to project forward the PSE.

Our census projections reflect males in urban areas of Guayas Province, where Guayaquil is located, based on figures from the Ecuadorian Statistics and Census Institute.

Projection could be improved by using census data specifically for males aged 15–49 years in Guayaquil, but such data were not available.

METHODS

USE THIS METHOD IF:

A	The <u>percent</u> of the KP or priority population estimated by the initial PSE is available. A census denominator for the desired year of implementation ("current year") is available.
B	Only the <u>number</u> of individuals estimated by the initial PSE is available and... Census denominators are available for the year when the PSE was conducted <u>and</u> for the current year. Typically, the denominator is either the male or female population aged 15 to 49 years.
C	Annual <u>growth rates</u> are available, starting from the year when the PSE was conducted to the year just before the current year (e.g., last year).

Enter the method to be used:

B

METHOD B

Derive the population percentage, then apply it to a current census denominator.	Estimate	Lower limit	Upper limit
Initial PSE as an absolute number of individuals	2,410	1,720	3,556
Census denominator for the general population in the year when the initial PSE was conducted	3,564,222	3,564,222	3,564,222
Census denominator for general population for the current year	3,616,285	3,616,285	3,616,285
Projected PSE	2,445	1,745	3,608

The projection, 2445 (95% CI 1745–3608) transgender women, is copied automatically to the Estimates Summary table. We document the details of the data source on the same row.

STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	2,445	1,745	3,608	National Institute of Statistics and Census (INEC) online population projections by province	Urban areas of Guayas province	Projections for 2017 and 2018 were used to determine the growth rate	General population	All ages
---	-------	-------	-------	--	--------------------------------	--	--------------------	----------

Step 4: Proportion HIV-negative

For the proportion of transgender women who are HIV-negative, we use the HIV prevalence estimate from the 2017 BBS. With an estimated prevalence of 33.2%, the proportion that is HIV-negative is 66.8%.

STEP 4. PROPORTION HIV-NEGATIVE	66.8%	53.5%	77.5%	RDS IBBS survey of TW	Guayaquil	2017	Persons assigned male sex at birth who describe their gender as female	≥ 18 years
--	-------	-------	-------	-----------------------	-----------	------	--	------------

Step 5: Subgroup inflation factor

There is no need to apply an inflation factor because, in theory, the PSE reflects the entire transgender women population in Guayaquil. More accurately, it represents all transgender women who could be reached through RDS and by unique objects, and who were willing to participate in both of these activities. Because we are not applying a subgroup inflator, we enter 100% for the estimates in Step 5.



STEP 5. SUBGROUP INFLATION FACTOR	100.0%	100.0%	100.0%	No inflator applied				
--	--------	--------	--------	---------------------	--	--	--	--



Step 6: Risk proportion

Step 6A: Choose representative data source

To estimate the proportion at substantial risk, we draw on the 2017 BBS of transgender women.

Step 6B: Define substantial risk of HIV infection

We use the risk factors approach to define the risk criteria. We draw on four established risk factors available in the BBS data: sexually transmitted infection; HIV status of recent anal male sex partners; number of recent anal sex partners; and unprotected anal intercourse. We defined two levels of risk based on these risk factors:

- ▶ High-risk: transgender women who tested positive for syphilis or self-reported sexually transmitted infection symptoms in the past 12 months.
- ▶ Medium-risk: other transgender women:
 - ▶ whose last anal sex episode was condomless and with a partner who they believed had HIV; or
 - ▶ who had 2 or more male anal sex partners and engaged in unprotected anal intercourse with any male partner in the past 12 months.

These definitions are not derived from any model, as in the exposures approach. They are subjective but based on risk factors identified in the literature. Many other potential categorization schemes are possible.

We fill out Section A of the Risk Proportion sheet to document the population included in the survey.

SECTION A. WHAT SURVEY DATA ARE BEING USED?

Fill in details about the survey that will be used to estimate risk behaviors.

Population surveyed:	Transwomen
Operational definition in survey:	Persons assigned male sex at birth who describe their gender as female
Age range or eligibility criteria	18 and older
Geographic area:	Guayaquil, Ecuador
Sampling method:	Convenience sample (Intended as RDS but did not reach sample size)
Year of survey:	2017

Step 6C: Classify risk level of survey participants

We used responses to question items to create measures of sexually transmitted infection; most recent anal sex that was unprotected anal intercourse and with a male partner living with HIV; number of male anal sex partners in the past 12 months; and unprotected anal intercourse with any partner in the past 12 months.

Before we classify participants into different risk levels, we look for skip patterns in the questionnaire that could cause problems in the three risk factor measures. This leads us to replace missing values with 0 in some cases. For example, some participants reported no stable or casual partners in the past 12 months and so were not asked questions about whether the most recent stable or casual partner was HIV-positive; we assign them a value of 0 for our people living with HIV measure. There were similar issues with the other measures.

We create a categorical variable, *risk_level*, coded 1 or 2 for the respective risk categories, being careful about missing values: we code *risk_level* as missing if the risk level is unclear due to non-response, so participants with incomplete data are excluded from the estimates.

RISK FACTORS APPROACH

SECTION D. HOW WERE THE RISK CATEGORIES DEFINED? WHAT PART OF THE POPULATION IS IN EACH CATEGORY?

Describe the risk criteria and the measures used to define risk. Three categories are listed. However, you may use one or two categories or as many as needed.

	RISK CATEGORY 1	RISK CATEGORY 2	RISK CATEGORY 3
How was this risk category defined?	Syphilis or STI symptoms in the past 12 months	Not in risk category 1 Last sex was condomless and with a PLWH or ≥ 2 male partners and any UAI in past 12 months	
What percent of all Transwomen are in this risk category?	24.30%	40.80%	
Lower limit	16.40%	31.20%	
Upper limit	33.70%	50.90%	

Copy the risk proportions above onto the PrEP Estimates worksheet (Step 6).

SECTION E. WHAT MEASURES WERE USED TO DEFINE THE RISK CATEGORIES?

Describe any measures utilized to define the risk categories above.

	Name of measure	How was this measure defined?
Measure 1	STI	Test positive for syphilis; or past 12-month STI symptoms, including: genital secretion; ulcers, sores, or warts in genital or anal area; and pain or burning when urinating.
Measure 2	Number of male partners	Number of male and TW sex partners in past 12 months
Measure 3	Any UAI	"Sometimes" or "never" (vs. "always") used a condom with any partner type over the past 12 months
Measure 4	UAI_PLWH	At last anal sex with a stable male partner, no condom was used and the partner had HIV; or at last sex with a casual male partner, no condom was used and the partner had HIV.
Measure 5		

In Sections D and E of the Risk Proportion sheet, we document how the risk factors and risk categories were defined.

Step 6D: Estimate proportion at risk among HIV-negative subjects

With the variable *risk_level* added to the dataset, we estimate the proportion at risk using the *ci* proportion command in Stata. We do not adjust for RDS because we are treating the data as a convenience sample given the small sample size.

Due to small sample size, we do not subset to participants who tested negative for HIV. We assume risk behaviour is similar regardless of HIV status.

We enter the estimates of 24.3% at risk level 1 and 40.8% at risk level 2 in Section E along with the 95% CIs. As in Step 4 (proportion who are HIV-negative), the CIs probably underestimate the uncertainty of the estimates due to convenience sampling.

The estimated risk proportions are copied automatically to the Estimates Summary table on the PrEP Estimates sheet. The estimated number in each risk category updates automatically.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data age range. Then compare the scope of the data sources used with the intended scope of the

	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	2,410	1,720	3,556
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	2,445	1,745	3,608
STEP 4. PROPORTION HIV-NEGATIVE	66.8%	53.5%	77.5%
STEP 5. SUBGROUP INFLATION FACTOR	100.0%	100.0%	100.0%
STEP 6. PROPORTION AT RISK			
ESTIMATED NUMBER AT RISK	0	0	0

ESTIMATES BY RISK CATEGORY

Fill in the box below only if using the Risk Factors Approach to define risk with more than one risk category.

	Estimate	Lower limit	Upper limit
Risk proportions:			
Risk Category 1	24.3%	16.4%	33.7%
Risk Category 2	40.8%	31.2%	50.9%
Risk Category 3			
Estimated numbers at risk:			
Risk Category 1	397	178	616
Risk Category 2	666	343	990
Risk Category 3	0	0	0

The estimates suggest that as of 2018, there were approximately 397 (uncertainty range 178–616) transgender women in Guayaquil who met our category 1 risk criteria and 666 (uncertainty range 343–990) who met our category 2 risk criteria and who were negative for HIV and aged 18 years or over.

Example 5: People who inject drugs—risk factors approach

This example estimates the number of people at risk who inject drugs in South Africa, as of 2017. Survey data are from an BBS that used convenience sampling and included males and females who inject drugs.

Step 1: Define the population

At the top of the PrEP Estimates sheet, we enter an operational definition of the population. Our aim is to estimate the number at risk among all males and females who currently inject drugs. We restrict our definition to people aged 15 years and over who frequent sites known for injection drug use, due to the venue-based PSE available.

STEP 1. DEFINE THE GEOGRAPHIC AREA AND POPULATION FOR TARGET-SETTING

Indicate the intended scope of the estimates:

Population:	PWID
Operational definition of population:	Males and females who inject drugs and frequent known IDU venues
Age range	15 and older
Geographic area:	National
Desired year of the estimate	2017

As we complete the remaining steps, we will be alert to possible misalignment of the data sources.

Step 2: Initial PSE

Moving down the Estimates Summary table, we enter details of the initial PSE. For people who inject drugs in South Africa, there was 1 PSE conducted using programmatic mapping nationally in 2015, with an estimate of 41 374–44 135 males and 31 489–34 402 females who inject drugs.

To get an estimate for males and females combined, we sum the upper and lower limits, leading to a range of 72 863–78 537, and set the main estimate as the midpoint ($= [72\ 863 + 78\ 537]/2$).

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data sources used in terms of geography, time period, definition of the population, and age range. Then compare the scope of the data sources used with the intended scope of the estimates (defined in Step 1) in order to identify any alignment issues.

	ENTER ESTIMATES			DOCUMENT THE DATA SOURCE FOR EACH ESTIMATE. IDENTIFY ANY ALIGNMENT ISSUES.				
	Estimate	Lower limit	Upper limit	Compare representativeness of each data source with the intended scope of the estimates (from Step 1)				
				What is the source of this estimate?	Geographic area	Year(s) when data were collected	How the population was defined	Age range
STEP 2. INITIAL PSE	75,700	72,863	78,537	PMSE: 41,374-44,135 males 31,489-34,402 females Figures were males and females were summed. Estimate is the midpoint of the range.	National	2015	Males and females who injected drugs in the past 12 months	≥ 15 years

Step 3: Project the PSE

Because a year has passed since the initial PSE was conducted, we project the PSE forward to the desired year of implementation, 2017, using a projection calculator (Project PSE sheet). We do not use method A because a population percentage is required, and the one available in the programmatic mapping size estimate was not based on the year it was conducted, as census data were not available for the same year, 2015.

Instead, we use method B. We enter the total size of the population (males and females) aged 15–49 years in 2015 and the year of the desired PrEP implementation, 2017. The spreadsheet uses these starting and ending population counts to calculate a growth rate and applies the rate to the initial PSE. We obtained the population counts from the United Nations Population Division.

METHODS

USE THIS METHOD IF:

The percent of the KP or priority population estimated by the initial PSE is available. A census denominator for the desired year of implementation ("current year") is available.

A

Only the number of individuals estimated by the initial PSE is available and...

B

Census denominators are available for the year when the PSE was conducted and for the current year. Typically, the denominator is either the male or female population aged 15 to 49 years.

C

Annual growth rates are available, starting from the year when the PSE was conducted to the year just before the current year (e.g., last year).

Enter the method to be used:

B

METHOD B

Derive the population percentage, then apply it to a current census denominator.	Estimate	Lower limit	Upper limit
Initial PSE as an absolute number of individuals	75,700	72,863	78,537
Census denominator for the general population in the year when the initial PSE was conducted	30,169,772	30,169,772	30,169,772
Census denominator for general population for the current year	30,920,940	30,920,940	30,920,940
Projected PSE	77,585	74,677	80,492

The projection, 77 585 (uncertainty range 74 677–80 492) people who inject drugs, is copied automatically to the Estimates Summary table. We document the details of the data source on the same row.

STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	77,585	74,677	80,492	UN Pop Division census denominators for males+females nationally	National	2015 & 2017	General population	15-49 years
---	--------	--------	--------	--	----------	-------------	--------------------	-------------

We could have used method C instead by downloading average annual growth rates for the years 2015 and 2016 for urban areas of the country from the World Bank.

By using method B, we can use a growth rate specifically for people aged 15–49 years but we cannot subset to urban areas. With method C, we can subset to urban areas but not to a specific age group. The local team should weigh the advantages and disadvantages of each approach in the local context.

Step 4: Proportion HIV-negative

We estimate the proportion of our target population, people who inject drugs, that is HIV-negative from an BBS survey. The survey was among males and females who inject drugs in five cities in South Africa. Even though the BBS was based on a convenience sample, we calculated a 95% CI in statistical software. In reality, there is probably more uncertainty in this estimate than the CI reflects. The range could be widened to account for this, but we do not do so here.

STEP 4. PROPORTION HIV-NEGATIVE	85.8%	82.2%	88.9%	Convenience sample IBBS (overall estimate for combined male and female sample) Limits were calculated as a 95% CI assuming a random sample.	Five cities in Gauteng (Centurion, Pretoria and Johannesburg), KwaZulu-Natal (Durban), and Western Cape Province (Cape Town)	2013	Males and females who had ever injected an illegal drug	≥ 18 years
--	-------	-------	-------	---	--	------	---	------------

Step 5: Subgroup inflation factor

Although the programmatic mapping size estimate was venue-based, there are no data on the percentage of people who inject drugs and frequent known injection sites that could be used as an inflator. Therefore, the final estimate is limited to venue-based people who inject drugs. In Step 5, we enter 100% (no inflation).

STEP 5. SUBGROUP INFLATION FACTOR	100.0%	100.0%	100.0%	No data				
--	--------	--------	--------	---------	--	--	--	--

Step 6: Risk proportion

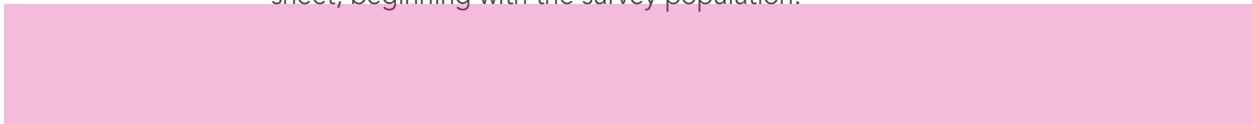
Step 6A: Choose representative data source

We select the most representative source of risk behaviours among people who inject drugs in the BBS mentioned previously.

Step 6B: Define substantial risk of HIV infection

We use the risk factors approach to define the risk criteria. We define survey participants as at risk if they either reported sharing a needle most or all times they inject drugs, or reported sexually transmitted infection symptoms in the past 12 months.

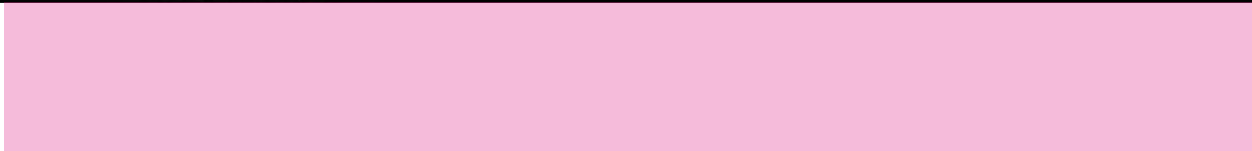
We will estimate the proportion at risk by conducting a new analysis of our dataset (as opposed to relying on a report), so we document this step on the Risk Proportion sheet, beginning with the survey population:



SECTION A. WHAT SURVEY DATA ARE BEING USED?

Fill in details about the survey that will be used to estimate risk behaviors.

Population surveyed:	PWID
Operational definition in survey:	Males and females who ever injected an illegal drug
Age range or eligibility criteria	18 and older
Geographic area:	5 cities in Gauteng (Centurion, Pretoria and Johannesburg), KwaZulu-Natal (Durban), and Western Cape Province (Cape Town)
Sampling method:	Convenience sample
Year of survey:	2013



Step 6C: Classify risk level of survey participants

We used responses to question items to create measures of sexually transmitted infection symptoms and needle-sharing.

Before we classify participants into different risk levels, we look for skip patterns in the questionnaire that could cause problems in the two risk factor measures. This leads us to replace missing values with 0 in some cases. For example, some participants reported no injecting in the past 12 months and so were not asked questions about needle-sharing. We assign them a value of 0 for our needle-sharing measure. There were similar issues with the other measure.

We create a dichotomous variable, *risk_level*, coded 1 if participants meet the risk criteria and 0 if they do not, being careful about missing values: we code *risk_level* as missing if the risk level is unclear due to non-response so participants with incomplete data are excluded from the estimates.

In Sections D and E of the Risk Proportion sheet, we document how the risk factors and risk categories are defined.

RISK FACTORS APPROACH

SECTION D. HOW WERE THE RISK CATEGORIES DEFINED? WHAT PART OF THE POPULATION IS IN EACH CATEGORY?

Describe the risk criteria and the measures used to define risk. Three categories are listed. However, you may use one or two categories or as many as needed.

	RISK CATEGORY 1	RISK CATEGORY 2	RISK CATEGORY 3
How was this risk category defined?	STI symptoms in the past 12 months or shares a needle always/most times		
What percent of all PWID are in this risk category?	67.50%		
Lower limit	62.60%		
Upper limit	72.20%		

Copy the risk proportions above onto the PrEP Estimates worksheet (Step 6).

SECTION E. WHAT MEASURES WERE USED TO DEFINE THE RISK CATEGORIES?

Describe any measures utilized to define the risk categories above.

	Name of measure	How was this measure defined?
Measure 1	sti symptoms	Self-report of STI symptoms in the past 12 months
Measure 2	sharing	Reported sharing a needle always or most times when injects drugs
Measure 3		
Measure 4		
Measure 5		

Step 6D: Estimate proportion at risk among HIV-negative subjects

With the variable *risk_level* added to the dataset, we estimate the proportion at risk using the *ci* proportion command in Stata. We run the estimate, subsetting the data to participants who tested negative for HIV.

We enter the estimate of 67.5% in Section E along with the 95% CI. As in Step 4 (proportion who were HIV-negative), the CI probably underestimates the uncertainty of the estimate because of convenience sampling.

These estimates are copied automatically to the Estimates Summary table on the PrEP Estimates sheet. The estimated number at risk in each risk category updates automatically.

The estimates suggest that as of 2017, there were approximately 44 923 (95% CI 40 908–48 937) people aged 15 years and over who inject drugs who were HIV-negative, frequented known injection sites in South Africa, and met the risk criteria.

Example 6: Adolescent girls and young women—risk factors approach

This example estimates the number of adolescent girls and young women in Lesotho in 2018 who fall into three different risk categories:

- ▶ High risk: lives with a partner living with HIV.
- ▶ Medium risk: had a sexually transmitted infection or sexually transmitted infection symptoms in the past year.
- ▶ Low risk: had multiple sex partners and had condomless sex with any of the last three sex partners in the past year.

Survey data are from a demographic and health survey (DHS) conducted in 2014. The DHS included HIV testing of women and a subsample of their cohabitating male partners, which is helpful when estimating the risk proportions among HIV-negative survey participants and identifying HIV-negative adolescent girls and young women with partners living with HIV.

Step 1: Define the population

At the top of the PrEP Estimates sheet, we enter an operational definition of the population as females aged 15–24 years nationally. This is the intended scope of the estimates.

STEP 1. DEFINE THE GEOGRAPHIC AREA AND POPULATION FOR TARGET-SETTING

Indicate the intended scope of the estimates:

Population: Adolescent girls and young women
Operational definition of population: Females aged 15 to 24 years
Age range 15 to 24 years
Geographic area: Lesotho
Desired year of the estimate 2018

As we complete the remaining steps, we will be alert to possible misalignment of the data sources.

Step 2: Initial PSE

We draw on the United Nations Population Division estimates of the number of females aged 15–24 years residing in Lesotho in 2018.³

We enter the resulting estimate of 241 837 females as the initial PSE on the row for Step 2, along with details of the data source. We enter the same estimate for the upper and lower limits. Uncertainty limits are not available for these population estimates.

ESTIMATES SUMMARY: CALCULATE THE ESTIMATES AND DOCUMENT DATA SOURCES

Enter the point estimate and limits for each step. On the right, describe the scope of the data sources used in terms of geography, time period, definition of the population, and age range. Then compare the scope of the data sources used with the intended scope of the estimates (defined in Step 1) in order to identify any alignment issues.

	ENTER ESTIMATES			DOCUMENT THE DATA SOURCE FOR EACH ESTIMATE. IDENTIFY ANY ALIGNMENT ISSUES.				
	Estimate	Lower limit	Upper limit	Compare representativeness of each data source with the intended scope of the estimates (from Step 1)				
				What is the source of this estimate?	Geographic area	Year(s) when data were collected	How the population was defined	Age range
STEP 2. INITIAL PSE	241,837	241,837	241,837	United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, custom data acquired via website: https://population.un.org/wpp/DataQuery	Lesotho	n/a (UN projection based on Census data)	Females ages 15 to 24 years	15-24 years

Step 3: Project the PSE

We do not need to project the PSE because it is for the same year for which we want the estimate of the number at risk, 2018. Since we are not projecting the PSE, we enter the initial PSE into the second row of the Estimates Summary table.

STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	241,837	241,837	241,837	No projection applied	
--	---------	---------	---------	-----------------------	--

³ See <https://population.un.org/wpp/Dataquery>. Select Population by age and sex, and refine the query by selecting Lesotho, year 2018, Female, and age groups 15–19 years and 20–24 years.

Step 4: Proportion HIV-negative

According to the DHS 2014 Lesotho report, the HIV prevalence among females aged 15–24 years nationally is 13.1% (95% CI 10.8–15.3%).

The proportion HIV-negative is 100% minus the prevalence, or 86.9%. The lower limit of the proportion HIV-negative is 100% minus the upper limit of the HIV prevalence. The upper limit of the proportion HIV-negative is 100% minus the lower limit of the HIV prevalence.

STEP 4. PROPORTION HIV-NEGATIVE	86.9%	84.7%	89.2%	Point estimate is from 'DHS Report Table 13.3, totals row (page 249). Limits are from DHS Report Table B.2. (Sampling Errors), bottom row (page 337) and reflect a 95% confidence interval.	Lesotho	2014	Females ages 15 to 24 years in households	15-24 years
--	-------	-------	-------	---	---------	------	---	-------------

Step 5: Subgroup inflation factor

A subgroup inflation factor is not needed because the PSE reflects the same population in which we are interested. We enter 100% in Step 5 of the Estimates Summary table.

STEP 5. SUBGROUP INFLATION FACTOR	100.0%	100.0%	100.0%	No inflator applied	
--	--------	--------	--------	---------------------	--

Step 6: Risk proportion

Step 6A: Choose representative data source

We used data from the 2014 Lesotho DHS to estimate the proportion of HIV-negative females aged 15–24 years who are at substantial risk for HIV infection. We chose the DHS because it provides a nationally representative sample and included HIV testing for all women and a subsample of male partners living in the same households, which is useful for estimating the proportion HIV-negative and the proportion who are HIV-negative with HIV-positive partners. The DHS also collected survey data on sexually transmitted infection and sexual behaviours.

The DHS is limited to the population accessible in households, but our intent is to estimate all adolescent girls and young women at risk. The risk proportion estimated based on the DHS will not reflect females who live outside households, such as in military barracks, prisons, hotels, brothels, student dormitories, and other organizational or institutional housing. Such biases should be discussed when using the estimates.

Step 6B: Define substantial risk of HIV infection

We use the risk factors approach to define the risk criteria. We review the risk factors shown to be related to incident HIV infection among adolescent girls and young women in the recent literature. (See the document *PrEP Target-setting for Key and High-priority Populations: Technical Materials*.) A number of proximal risk factors are shown, including:

- ▶ Number of recent sex partners.
- ▶ Number of lifetime sex partners.
- ▶ Partner living with HIV or newly diagnosed with HIV.
- ▶ Partner of unknown HIV status.
- ▶ Sexually transmitted infection history or treatment.

We review the DHS questionnaire to determine which of these risk factors can be assessed using the DHS data. In addition to the questionnaire, we can also review the Standard Recode Manual for DHS 6, which lists the variables in standardized form.⁴

We use the following three risk categories, drawing largely on the proximal risk shown in the technical materials annex. We include condomless sex as part of the definition, based on theory, even though it is not listed in the review findings:

- ▶ High risk: lives with a partner living with HIV.
- ▶ Medium risk: had a sexually transmitted infection or sexually transmitted infection symptoms in the past year.
- ▶ Low risk: had multiple sex partners and had condomless sex with any of the last three sex partners in the past year.

These risk categories are limited to some extent due to limitations of DHS data. For example, we would prefer not to restrict the high-risk category to females who live with an partner living with HIV, but we are limited by the fact that the DHS only tested

⁴ See https://www.dhsprogram.com/pubs/pdf/DHSG4/Recode6_DHS_22March2013_DHSG4.pdf.

cohabitating partners. Similarly, the DHS asked about condom use only with the three most recent sex partners.

Many other risk definitions are possible. For example, the definition above does not include some of the proximal risk factors (e.g. lifetime number of partners) for which the literature review in the document *PrEP Target-setting for Key and High-priority Populations: Technical Materials* found evidence, or any of the distal risk factors, although they could be assessed with the DHS data.

To document the risk definition, we describe the survey population in Section A of the Risk Proportion sheet.

SECTION A. WHAT SURVEY DATA ARE BEING USED?

Fill in details about the survey that will be used to estimate risk behaviors.

Population surveyed:	Demographic health survey (DHS): general population accessible at households with HIV testing of males and females
Operational definition in survey:	Females and males in households
Age range or eligibility criteria	15 to 49 years
Geographic area:	Lesotho
Sampling method:	Multi-stage probability sample of individuals within households within enumeration areas
Year of survey:	2014

Step 6C: Classify risk level of survey participants

To assess the risk criteria using DHS data, we create a dataset of all female participants that includes their HIV test results and the HIV test results of any cohabitating partners who were tested. To create the file:

- ▶ Merge the women’s data file with the HIV test results data file.
- ▶ Merge the men’s data file with the HIV test results data file.
- ▶ Merge the women’s and men’s data files, linking on the line number of each woman’s primary partner, if any.⁵

We define variables representing the three criteria. For sexually transmitted infections, we include self-report of having a sexually transmitted infection and of symptoms (genital sore, ulcer, discharge). For condomless sex, we cannot use the measure on consistent condom use because of high missingness (24–54% of responses are incomplete) and use condom use at last sex instead.

Missing data and sampling weights

The first risk category is missing for 64% of female participants who reported living with a partner because their partner was not tested. The measure was complete for all participants who did not live with a partner, because it was set to 0. This leaves out a disproportionately large share of cohabitating female participants in a way that could bias the risk proportion.

⁵ For guidance on merging DHS datasets, see <https://dhsprogram.com/data/Merging-Datasets.cfm> and the entry Merging DHS data in Stata on the DHS Program User Forum at <https://userforum.dhsprogram.com>.

In a way, the 36% of cohabitating females with complete data must “represent” the other cohabitating females with incomplete data. We therefore increase the sampling weights of the cohabitating females who had complete data on this measure by the amount of sampling weight lost due to excluding cohabitating women who had incomplete data.

Checking skip patterns

Before we classify participants into different risk levels, we look for skip patterns in the questionnaire that could cause problems in the three risk factor measures. We assign 0 to participants who did not have a cohabitating partner (as described above) for the first risk category. For the third measure on sexual behaviour, we are careful not to exclude participants who did not report condom use for the first, second or third partner if they did not have as many partners.

We create a categorical variable, *risk_level*, coded 1, 2 or 3 for the respective risk categories, being careful about missing values. We code *risk_level* as missing if any of the three risk category variables is missing so participants with incomplete data are excluded from the estimates.

In Sections D and E of the Risk Proportion sheet, we document how the risk factors and risk categories are defined.

RISK FACTORS APPROACH

SECTION D. HOW WERE THE RISK CATEGORIES DEFINED? WHAT PART OF THE POPULATION IS IN EACH CATEGORY?

Describe the risk criteria and the measures used to define risk. Three categories are listed. However, you may use one or two categories or as many as needed.

	RISK CATEGORY 1	RISK CATEGORY 2	RISK CATEGORY 3
How was this risk category defined?	Cohabitating partner living with HIV	STI or STI symptoms in past year	Multiple sex partners and condomless sex with any of the last 3 sex partners in past year
What percent of all Demographic health survey (DHS): general population accessible at households with HIV testing of males and females are in this risk category?	1.68%	11.35%	1.51%
Lower limit	0.77%	9.23%	0.90%
Upper limit	3.61%	13.89%	2.50%

Copy the risk proportions above onto the PrEP Estimates worksheet (Step 6).

SECTION E. WHAT MEASURES WERE USED TO DEFINE THE RISK CATEGORIES?

Describe any measures utilized to define the risk categories above.

	Name of measure	How was this measure defined?
Measure 1	cohab_hiv	Had a cohabitating partner living at home at the time of the survey with HIV
Measure 2	sti	Had STI or STI symptoms (genital sore/ulcer/discharge) in past 12 months
Measure 3	sex_behav	Had two or more sex partners in past 12 months and did not use a condom at last sex with any of last 3 partners in the past 12 months
Measure 4		
Measure 5		

Step 6D: Estimate proportion at risk among HIV-negative subjects

With the variable *risk_level* added to the dataset, we estimate the proportions of adolescent girls and young women in each of the three risk categories at risk.

DHS data must be weighted and adjusted for clustering when estimating proportions and CIs. We apply the women's weight variable (*v005*) and adjust for clustering using the cluster (*v001*) as the primary sampling unit and the stratum identifier (*v023*).⁶

In Stata, we use the *svyset* and *svy, subpop:tab* commands to produce the adjusted estimates, restricted to the subpopulation of HIV-negative females aged 15–24 years.

We enter the estimates of 1.6% adolescent girls and young people at high risk, 11.35% at medium risk, and 1.51% at low risk in Section E along with the 95% CIs. We enter the same estimates into the Estimates by Risk Category section of the PrEP Estimates sheet. The estimated number in each risk category updates automatically.

The estimates suggest that as of 2018, there were 3510 adolescent girls and young women in Lesotho at high risk, 23 853 at medium risk, and 3173 at low risk, according to this risk definition.

⁶ For tutorials on weighting DHS data, see <https://blog.dhsprogram.com/sampling-weighting-at-dhs/>.

	ENTER ESTIMATES		
	Estimate	Lower limit	Upper limit
STEP 2. INITIAL PSE	241,837	241,837	241,837
STEP 3. PSE PROJECTED TO CURRENT YEAR (copied from Project PSE tab)	241,837	241,837	241,837
STEP 4. PROPORTION HIV-NEGATIVE	86.9%	84.7%	89.2%
STEP 5. SUBGROUP INFLATION FACTOR	100.0%	100.0%	100.0%
STEP 6. PROPORTION AT RISK			
ESTIMATED NUMBER AT RISK	0	0	0

ESTIMATES BY RISK CATEGORY

Fill in the box below only if using the Risk Factors Approach to define risk with more than one risk category.

	Estimate	Lower limit	Upper limit
Risk proportions:			
Risk Category 1	1.7%	0.8%	3.6%
Risk Category 2	11.4%	9.2%	13.9%
Risk Category 3	1.5%	0.9%	2.5%
Estimated numbers at risk:			
Risk Category 1	3,510	524	6,495
Risk Category 2	23,853	18,917	28,788
Risk Category 3	3,173	1,490	4,857



UNAIDS
Joint United Nations
Programme on HIV/AIDS

20 Avenue Appia
1211 Geneva 27
Switzerland

+41 22 791 3666

unaids.org