



**Overview of making estimates of  
HIV/AIDS and its impact in countries  
with low-level or concentrated epidemics:  
The Workbook Method**

The models and methodology of the UNAIDS/WHO  
approach to estimating and projecting national HIV/AIDS  
epidemics

The UNAIDS Reference Group on Estimates, Models and  
Projections

June 2003

---

# Table of Contents:

<b><i>I. Overview of making estimates of HIV/AIDS and its impact in countries with low-level or concentrated epidemics</i></b>	<b>3</b>
<b><i>II. Point-Prevalence Workbook</i></b>	<b>4</b>
<b>A. Overview of the process</b>	<b>4</b>
<b>B. Structure of the Point-Prevalence Workbook</b>	<b>5</b>
<b>C. Basic steps in using the Point Prevalence Workbook to make an estimate of national prevalence</b>	<b>7</b>
1. <i>Defining and entering regions.</i>	7
2. <i>Entering population information.</i>	8
3. <i>Selecting and inserting groups at higher risk (PHR).</i>	8
4. <i>Estimating the population size for the populations at higher risk (PHRs).</i>	9
5. <i>Estimating prevalence in the populations at higher risk (PHRs).</i>	10
6. <i>Selecting and inserting population at lower risk (PLR).</i>	10
7. <i>Selecting which approach to use in estimating prevalence among those at lower risk.</i>	12
<b>D. Results and outputs</b>	<b>12</b>
<b>E. Checks and reminders after making the initial estimate</b>	<b>13</b>
<b><i>III. Projection Workbook</i></b>	<b>15</b>
<b>A. Overview of the process</b>	<b>15</b>
<b>B. Structure of the Projection Workbook</b>	<b>15</b>
<b>C. Basic steps in using the Projection Workbook to develop an epidemic curve and short-term projection of the epidemic</b>	<b>18</b>
1. <i>Base year of estimate and projection.</i>	18
2. <i>Adult population.</i>	18
3. <i>Populations at higher risk.</i>	19
4. <i>Population sizes.</i>	19
5. <i>Prevalence in populations at higher risk</i>	20
6. <i>Growth in prevalence among populations at higher risk.</i>	20
7. <i>Population at lower risk.</i>	21
8. <i>Entering previous estimates for the curve fitting process.</i>	22
9. <i>Setting the curve fitting parameters, and producing the best-fitting curve.</i>	23
10. <i>Running the curve fitter.</i>	24
<b>D. Results and outputs</b>	<b>25</b>
<b>E. Checks and reminders after making the initial epidemic curve and short-term projection</b>	<b>25</b>

# **I. Overview of making estimates of HIV/AIDS and its impact in countries with low-level or concentrated epidemics**

This manual will provide a brief overview of using spreadsheets to create and estimate and projections of HIV/AIDS in countries with a low-level or concentrated epidemic. It will present the major structure and then walk you through the steps required to make an estimate using the Point Prevalence Workbook and then do the same for making short-term projections using the Projections Workbook. The manual does not include an overview of the software Spectrum, as that is described elsewhere.

## **A three-step process**

1. Make a point prevalence estimate for current and some past years using the *Point Prevalence Workbook*
2. Make an epidemic curve for the national estimate of prevalence over time using the *Projections Workbook*. This could include making a short-term projections
3. Make estimates of incidence, mortality and impact of HIV/AIDS using *Spectrum*.

*Note:* You may also want to use *EPP* to derive separate epidemic curves for various groups at higher risk for which you have time series data on prevalence.

## II. Point-Prevalence Workbook

### A. Overview of the process

You must make several general decisions about how you will structure the approach to make a national estimate. The workbook is set up so that you can define epidemics differently in different countries and regions, and still use the workbook to make an estimate of prevalence. The key is to make the decisions first about the structure you would like to use in making the estimate of prevalence then based on those decisions define that structure in the workbook to make the estimate.

*First, you must decide about how many geographically distinct epidemics and estimates you will use in making the national estimate.* This decision should be based on a mix of political, practical and epidemiological considerations. In very large or geographically disparate countries, there should be estimates made for each region. Even in smaller countries, politically it may be helpful to have an estimate for a region or province. However, one must have data about population sizes of groups at higher risk of HIV/AIDS and prevalence at the regional levels for these estimates to be meaningful

*Second, within each region and at the national level, you must decide which groups or populations are exposed to HIV/AIDS.* The total estimate of prevalence will be the sum of the number of people living with HIV/AIDS in groups at higher and lower risk. There will be little or no spread among the low-risk populations, instead prevalence in the low-risk groups will be from interaction between low and high-risk populations (e.g. through sexual partners from the high risk population) or due to unsafe blood transfusion or medical injection practices. While some populations at risk exist in almost all countries (e.g., sex workers and their clients) others will be specific to a country or region. For each of these populations you will need an estimate of the size of population and HIV prevalence in the population.

*Third, you will have to decide how to handle estimating prevalence in the population at lower exposure to HIV/AIDS.* For this population we assume that they have no risk of exposure to HIV/AIDS except due to sexual interaction with their partner who is from a

group at higher risk. For example, these could be spouses of clients of sex workers or injecting drug users. Early on in epidemics, the number of people in this group who have HIV will be very small, as there will not be time enough for much spread via sex between the groups at higher risk and their sexual partners. Over time, the number of people with HIV/AIDS in this group can be much larger. Early in the epidemic, estimates of the number of people with HIV/AIDS in this group may best be calculated by directly estimating sexual transmission. Later on the epidemic, one may use prevalence in low risk populations such as in pregnant women as the basis for the estimate.

## **B. Structure of the Point-Prevalence Workbook**

Overall there are 14 sheets in the workbook that you can view. They are:

- ◆ Regional sheets

This spreadsheet is set up so that you can calculate prevalence for up to 35 regions. This is done by filling in information on any of the regional sheets (e.g., Region 1, Reg 2, etc.) and the single "Remaining Region" sheet. Each regional sheet has the same structure and is used to estimate the number of adults living with HIV/AIDS in that region. If you use less than 10 regions it will still work, but you must at least fill in the sheet "Remaining Regions" to produce a national estimate. Once the regional information has been added, the "National estimate" sheet will automatically sum up the regional sheets to produce a national estimate.

On these sheets, you also need to enter estimates of percent females in each group.

*Notes:* The regional sheet names can be renamed with the actual region name (e.g., "Region 1" may be renamed to "Delhi"). Also there are additional regional sheets that are linked to the national estimate sheet, but all but 10 have been hidden. If you need additional regional sheets, go to menu "Format" and under "Columns" select "Unhide". This will make additional regional sheets available to use.

2002	Population Sizes Estimates		Prevalence Estimates (%)		Estimates of People living with HIV/AIDS				Average PLVHA	Percent [%] female in risk group	Number of women Infected	Percent [%] of infected who are women
	Country Name	Low	High	Low	High	(Low Population x Low Prevalence)	(Low Population x High Prevalence)	(High Population x Low Prevalence)				
Total Adult population 15-49	2,500,000											
<b>1. Populations at higher risk (PHR)</b>												
IDU	0	0	0	0	0	0	0	0	0	0	0	0
MSM	0	0	0	0	0	0	0	0	0	0	0	0
Sex workers	0	0	0	0	0	0	0	0	0	0	0	0
Clients of sex workers	0	0	0	0	0	0	0	0	0	0	0	0
Optional HPR1	0	0	0	0	0	0	0	0	0	0	0	0
Optional HPR2	0	0	0	0	0	0	0	0	0	0	0	0
Optional HPR3	0	0	0	0	0	0	0	0	0	0	0	0
Optional HPR4	0	0	0	0	0	0	0	0	0	0	0	0
Sub-Total PHR	0	0							-		0	0
<b>2. Populations at lower risk (PLR) that are not already included in PHR</b>												
<b>a. Partners of high risk populations</b>												
Partners of IDU	0	0	0	0	0	0	0	0	0	0	0	0
Female partners of MSM	0	0	0	0	0	0	0	0	0	0	0	0
Partners of Clients of Sex workers	0	0	0	0	0	0	0	0	0	0	0	0
Optional LPR1	0	0	0	0	0	0	0	0	0	0	0	0
Optional LPR2	0	0	0	0	0	0	0	0	0	0	0	0
Optional LPR3	0	0	0	0	0	0	0	0	0	0	0	0
Sub-total Partners of high risk	0	0							0		0	0
<b>b. ANC data applied to low risk women</b>												
Urban female low risk pop	0	0							0		0	0
Rural female low risk pop	1,250,000	1,250,000			0.00%	0.00%			0		0	0

◆ "National Estimate" sheet

This sheet sums the inputs from the regional sheets to produce the national estimate, shown in the yellow cells.

For People at Higher Risk (PHR), all the inputs from the regional sheets are shown and used to calculate the national PHR sub-total.

For People at Lower Risk (PLR), all the inputs from the regional sheets that are selected for in cells G19 and G20, are used to calculate the regional sub-total and are included in the national estimate.

This sheet also has some consistency checks at the bottom of the estimates to see if your estimate is producing consistent results.

◆ "National Population" sheet

This sheet is where you include the national population (total and 15-49) that are from the UN Population Division (2002 revision). These can be replaced with other population estimates made by the country. For regions, you must provide all the population figures on the regional sheets.

- ◆ "Data Sheet"

Place the data you use to make the regional estimates in this sheet. Include data sources where available.

- ◆ "Notes"

This sheet has a brief description of the Point Prevalence Workbook.

**In the workbook, cells are color-coded based on function.**

- ◆ Cells that are colored **light blue** need your input.
- ◆ Cells that are colored **light green** are optional.
- ◆ Cells that are **light orange** have calculated values, based on what you put in the blue cells. Cells that are in light orange are also locked so that you will not accidentally change the formulas. However, if you wish to change the formula, just unprotect the sheet.
- ◆ Cells that are **yellow** contain the national estimate.

## **C. Basic steps in using the Point Prevalence Workbook to make an estimate of national prevalence**

### *1. Defining and entering regions.*

First, you need to decide the number and names of regions that will be used in making your national estimate. You can rename the sheets by highlighting the tab and then typing in the name of the region. Use as many Regional sheets as needed but always use the sheet named "Remaining Region" as the final regional sheet. For each region that you decide to use in making the national estimate you must do steps 2 through 6 described below. The combined estimates from each of the regions will then be the national estimate. Again, if you need additional sheets for more regions, they can be accessed by going to the "Format" menu and under "Columns" using "Unhide" to make more sheets appear.

## 2. Entering population information.

On the “National Population” sheet you must identify the year for which you are making the estimate in cell A12 and the national (15-49) population for that year in cell B12. The year included will appear on all sheets and the national population will be used to calculate the national prevalence. On each regional sheet used you must also include that region’s (15-49) population, and percent of the population that lives in urban areas in that region.

Age Group (15-49)	2000	2001	2002	2003	2004	2005
All ages						
Year	Population (15-49)					
2002	2,500,000					

## 3. Selecting and inserting groups at higher risk (PHR).

You must identify the groups at higher risk of infection that you will use to make your estimates. The worksheets are set up with four default groups:

- ◆ Injecting drug users (IDU),
- ◆ men who have sex with men (MSM),
- ◆ female sex workers, and
- ◆ clients of female sex workers.

If one group does not apply, you can simply set the population size to zero.

If you wish to add a group at higher risk, you can do it in one of the four optional groups that are in the spreadsheet. Optional group names that will be used can be added. Simply highlight the cell [e.g., Cell A11] and type in the new group name. If you do this on the sheet named "Region 1" this will also transfer those risk groups to all the other region sheets.

2002	Population Sizes Estimates		Prevalence Estimates (%)		Estimates of People living with HIV/AIDS				Average PLWHA	Percent (% female in risk group)	Number of women Infected	Percent infected are women
Regional Name	Low	High	Low	High	(Low Population x Low Prevalence)	(Low Population x High Prevalence)	(High Population x Low Prevalence)	(High Population x High Prevalence)				
Region Adult population (15-49)												
% Urban population												
<b>1. Populations at higher risk (PHR)</b>												
IDU					0	0	0	0	0			0
MSM					0	0	0	0	0			0
Sex workers					0	0	0	0	0			0
Clients of sex workers					0	0	0	0	0			0
Optional HR1					0	0	0	0	0			0
Optional HR2					0	0	0	0	0			0
Optional HR3					0	0	0	0	0			0
Optional HR4					0	0	0	0	0			0
<b>Sub-Total PHR</b>	<b>0</b>	<b>0</b>							<b>0</b>			<b>0</b>
<b>2. Populations at lower risk (PLR) that are not already included in PHR</b>												
					Please select one!							
					Select one:	PLR						
						ANC data						
<b>a. Partners of high risk populations</b>												
Partners of IDU					0	0	0	0	0			0
Female partners of MSM					0	0	0	0	0			0
Partners of Clients of Sex workers					0	0	0	0	0			0
Optional LR1					0	0	0	0	0			0
Optional LR2					0	0	0	0	0			0
Optional LR3					0	0	0	0	0			0
<b>Sub-Total Partners of high risk</b>	<b>0</b>	<b>0</b>							<b>0</b>			<b>0</b>
<b>b. ANC data applied to low risk women</b>												
Urban female low risk pop	0	0			0	0	0	0	0			0
Rural female low risk pop	0	0			0	0	0	0	0			0
<b>Sub-Total of low risk women</b>	<b>0</b>	<b>0</b>							<b>0</b>	<b>100.0%</b>		<b>0</b>

4. *Estimating the population size for the populations at higher risk (PHRs).*  
 Put in the low and high estimates of the population size for each of the groups at higher risk. For each group you will need an estimate of the size of the population [Cells B7 to B14 for low and C7 to C14 for high]. As population size estimates are difficult to make accurately, we suggest using both a low and a high estimate for each population. One could use a minimum and maximum value, or if one has a good estimate of population size, one could use a lower and upper confidence bound around that estimate.

*Note:* The low and high estimates of the population sizes (and also prevalence) should reflect the uncertainty around the estimates. Few countries have nationally

representative surveys to estimate size of the populations at higher risk; instead, programmes use various indirect measures to make the estimates. If one does have a national representative estimate then you can use the confidence interval to determine the range around the estimate. For populations where a nationally representative survey is not available, one should use larger ranges, with the absolute values used based on your judgment of the quality of the estimate.

*5. Estimating prevalence in the populations at higher risk (PHRs).*

For each population you will also need to have low and high estimates of prevalence for each group [Cells D7 to D14 for low, E7 to E14 for high]. Ideally the range used would be based on measure prevalence in each population, but often few data are available. One should also include the percentage of the populations at higher risk that are women [Cells K7 to K14]. This information will automatically calculate the population of low-risk women as well as an estimate of the female to male ratio in the overall estimate.

*Note:* Once you have entered the estimates of population size and prevalence for a group the estimated average number of PLWHA in that group will be automatically calculated. The five estimates of the number of people living with HIV/AIDS in that group will appear: Low-low the result of low population estimate times low prevalence estimate [Cell F7 to F14] through High-high: the product of high population size estimate time high prevalence estimate [I7 to I14]. The Low-high estimates [Cells G7 to G14] and High-low estimates [Cells H7 to H14] as well as the average value based on the four combinations [Cells J7 to J14] are also calculated.

*6. Selecting and inserting population at lower risk (PLR).*

There are two ways to estimate the number of people who are at low risk of infection except as regular sexual partners of members of the groups at higher risk (e.g., the wife of a man who has sex with men). One way is to estimate the number of sexual partners of people with high-risk behaviours and then estimate prevalence in this group. To make the estimate of the number of sexual partners you will have to have behavioural information about the high-risk groups (e.g., the percentage of men who have sex with men who are married or have regular female sex partners) [Cells B22

to B27 and C22 to C27]. Prevalence in these groups is usually not available. To estimate prevalence in these sexual partners, one can estimate prevalence by frequency of sex and probability of transmission during heterosexual sex. For example, if one assumes that sex occurs once a week and the transmission probability (male to female) is 0.2% per act, then 10% of the low risk sexual partners will be infected in a year. Using these types of calculations one can get a rough estimate of prevalence for these sexual partners of populations at higher risk.

2002	Population Sizes Estimates		Prevalence Estimates (%)		Estimates of People living with HIV/AIDS				Average PLWHA	Percent (%), female in risk group	Number of women Infected	Percent infected are w...
Regional Name	Low	High	Low	High	(Low Population x Low Prevalence)	(Low Population x High Prevalence)	(High Population x Low Prevalence)	(High Population x High Prevalence)				
<b>2. Populations at lower risk (PLR) that are not already included in PHR</b>					<b>Please select one!</b>							
					Select one:	PLR						
					ANC data							
<b>a. Partners of high risk populations</b>												
Partners of IDU					0	0	0	0	0		0	
Female partners of MSM					0	0	0	0	0		0	
Partners of Clients of Sex workers					0	0	0	0	0		0	
Optional LR1					0	0	0	0	0		0	
Optional LR2					0	0	0	0	0		0	
Optional LR3					0	0	0	0	0		0	
<b>Sub-total Partners of high risk</b>	<b>0</b>	<b>0</b>							<b>0</b>		<b>0</b>	
<b>b. ANC data applied to low risk women</b>												
Urban female low risk pop	0	0			0	0	0	0	0			
Rural female low risk pop	0	0			0	0	0	0	0			
<b>Sub-total of low risk women</b>	<b>0</b>	<b>0</b>							<b>0</b>	<b>100.0%</b>	<b>0</b>	
<b>Sub-Total PLR</b>									<b>0</b>		<b>0</b>	
<b>No Risk Population</b>	<b>0</b>	<b>0</b>										
<b>TOTALS</b>	<b>0</b>	<b>0</b>							<b>0</b>		<b>0</b>	
<b>Notes/Comments:</b>												

*Note:* These sexual partners should be assumed to have no other risk factors otherwise they will be included in the groups at higher risk estimates. If you wish to include this information in the final regional and country estimate, then G19 must be selected for that region. You must identify the groups at lower risk of infection that you will use to make your estimates. The usual groups are the low risk sexual partners of injecting drug users (IDU), female partners of men who have sex with men (MSM), and partners/wives of clients of female sex workers.

The second way to estimate the number of people in the low risk populations who have HIV/AIDS is to use prevalence data from pregnant women as the prevalence in

the populations at low risk. To use this approach one would use the number of women in the reproductive age (15-49) minus those women in the populations at high risk. The population size will be calculated automatically if you have included percentage of women in the high-risk groups and have entered the population size and percent of the population living in urban areas [Cells B3 and B4]. The population is the total number of women (15 to 49) minus women in the populations at higher risk [Cells B-C 31 to B-C32]. Then one would use as the prevalence rate in this group, the HIV prevalence found in urban and rural pregnant women.

*7. Selecting which approach to use in estimating prevalence among those at lower risk.*

It is good to use both approaches to calculating the number of people infected from the population at lower risk. However, to do the final estimate, you must select one method to be included. This you do by placing an “x” in one cell next to the choice [Cells G19 or G20]. You can change your choice and the estimate will be changed automatically to reflect the approach you have chosen. You may choose different options on different regional sheets. The national estimate sheet will calculate the national estimate using only those approaches selected for on the regional sheets.

#### D. Results and outputs

Once you have filled out each of the regional sheets, the sheet titled “National Estimate” will automatically be filled. This sheet has the same structure as the region sheets, but sums up the values from each of the region sheets. For example, the estimates of the population size of injecting drug users will be the sum of each of the regions. Likewise the estimate of the number of injecting drug users with HIV/AIDS is the sum of the estimates in all of the regions. The national estimates of prevalence in each group (low and high) are based on the sums of the number infected in each group divided by the population size of the group.

In addition to the cells that are the summary of the region sheets, the “National Estimate Sheet” has two additional sections. These are a National Estimate summary section [Cells A-B41 to A-B45] and a series of consistency checks [Cells A-E47 to A-E56]

which compare the values used in your national estimate to those in other countries with low-level or concentrated epidemics. The consistency checks will be described below.

2002	Population Sizes Estimates		Prevalence Estimates (%)		Estimates of People living with HIV/AIDS				Average PLVHA	Percent (%) female in risk group	Number of women Infected	Percent (%) of infected who are women
Country Name	Low	High	Low	High	(Low Population x Low Prevalence)	(Low Population x High Prevalence)	(High Population x Low Prevalence)	(High Population x High Prevalence)				
Rural female low risk pop	1,250,000	1,250,000	0.00%	0.00%	0	0	0	0	0		0	0
<b>Sub-total of low risk women</b>	<b>1,250,000</b>	<b>1,250,000</b>							<b>0</b>		<b>0</b>	<b>0</b>
<b>Sub-Total PLR</b>									<b>0</b>		<b>0</b>	<b>0</b>
<b>No Risk Population - NA</b>	<b>1,250,000</b>	<b>1,250,000</b>										
<b>TOTALS</b>	<b>1,250,000</b>	<b>1,250,000</b>							<b>0</b>		<b>0</b>	<b>0</b>
<b>National Estimates for year: 2002</b>			<b>Notes/Comments:</b>									
<b>Number of Adults (15-49) LWHA</b>	<b>0</b>											
<b>Adult Prevalence (15-49)</b>	<b>0.00%</b>											
<b>Number of Women (15-49) LWHA</b>	<b>0</b>											
<b>% of adults (15-49) who are women</b>	<b>0</b>											
<b>Consistency Checks</b>												
% of total population (15-49) who are IDUs	0.0%	<b>ok!</b>	While the extent of injecting drug use varies dramatically among countries few countries will have more than 0.7% of the adult (15-49) population who inject drugs.									
% of men (15-49) who are MSM	0.0%	<b>Unusually LOW value!</b>	Research has found that in most countries between 2% and 5% of men aged 15-49 have sex with other men.									
% of women (15-49) who are sex workers	0.0%	<b>Unusually LOW value!</b>	Few countries have good estimates of the number of sex workers. In Thailand the estimated number of sex workers is roughly 0.8% of the female population (15-49).									
% of men (15-49) clients of female sex workers	0.0%	<b>Unusually LOW value!</b>	Few countries have good estimates of the number of clients of sex workers. In Thailand the estimated number of clients is roughly between 5% and 20% of the male population (15-49).									
PLR to PHR ratio	0		Using this approach to estimating prevalence, the majority of people living with HIV/AIDS should be from your groups at higher risk. If the ratio is greater than .33 you may have over-estimated the prevalence in the low-risk population.									
HIV prevalence rate (%) in IDUs	0.0%	<b>ok!</b>										
HIV prevalence rate (%) in MSM	0.0%	<b>ok!</b>										
HIV prevalence rate (%) in sex workers	0.0%	<b>ok!</b>										
HIV prevalence rate (%) in clients of sex workers	0.0%	<b>ok!</b>										

### E. Checks and reminders after making the initial estimate

The series of consistency checks in the cells A-E47 to A-E56 compare values related to average population sizes of groups at high risk and estimates of prevalence in these groups to ranges found in other countries with low-level and concentrated epidemics. These values are not meant to be proscriptive; rather they simply give a comparison to average values. If a value you used in making your national estimate falls outside of this range, a warning is given (e.g., “The percentage of men who have sex with men is unusually low”). If the values are within the range, “OK” is written in the cell [C48 to C56].

Once you are satisfied with the national estimate for the year, you can then proceed to develop an epidemic curve and make short-term projection for the national epidemic. To do this, you will need at least three point-prevalence national estimates. For these estimates you can either use earlier estimates of national prevalence, or you can use the Point Prevalence Workbook to make an estimate for the earlier years. It is recommended to also include national point prevalence estimates from years that reflect both the early and middle stages of epidemic when developing the epidemic curve.

### **III. Projection Workbook**

#### **A. Overview of the process**

The projection workbook is based on the same general assumptions of the point prevalence workbook. That is, estimates and the short-term projection of the epidemic will assume that the primary spread of the epidemic will be in populations at higher risk (e.g., men who have sex with men, injecting drug users). In addition, there will also be limited spread to low-risk populations due to heterosexual transmission.

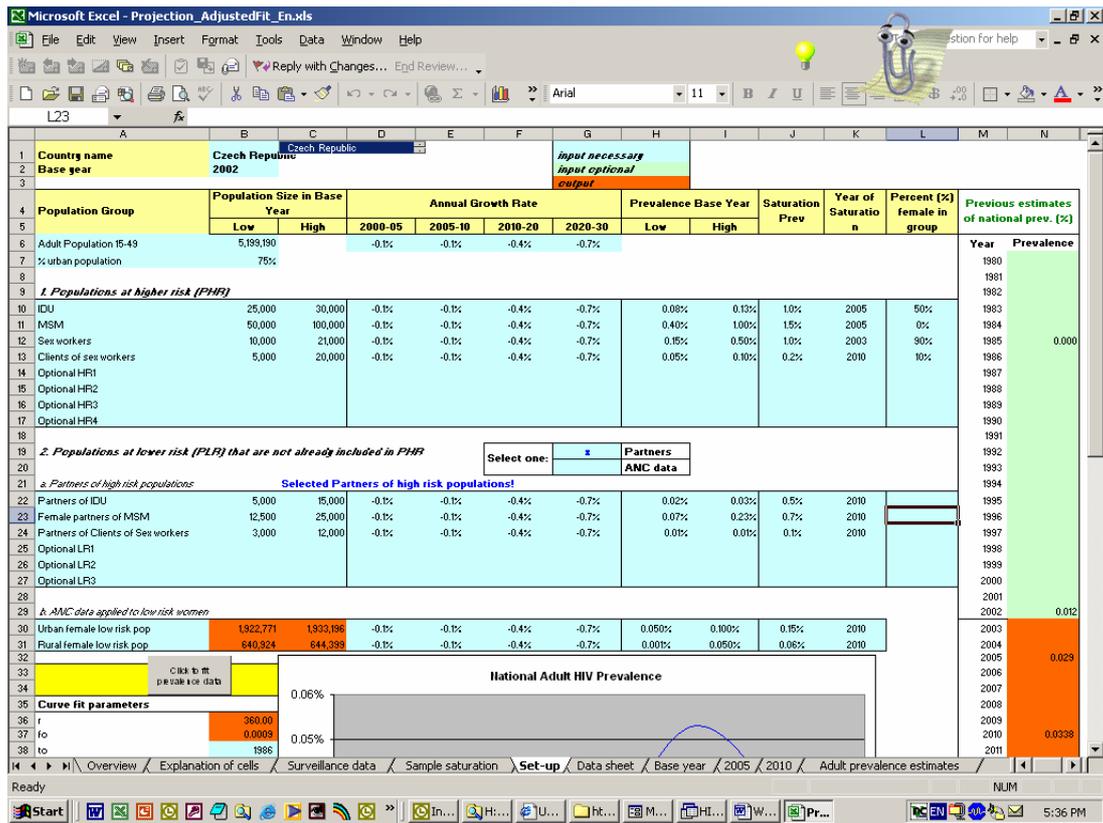
In addition to the point prevalence estimate for a single year, the workbook will also use point prevalence estimates and a curve-fitting approach to produce an epidemic curve for national prevalence. The workbook will also force you to make assumptions about the saturation level of the HIV/AIDS in populations at higher risk of infection to make short-term projections.

#### **B. Structure of the Projection Workbook**

##### **Process and flow of the model**

This projection model is made up of a "Set-up" sheet, yearly prevalence output sheets (for "Base Year", "2005" and "2010") and an "Adult prevalence estimates" output sheet. All data entry and assumptions are done in the "Set-up" sheet. Data and assumptions entered on this sheet are automatically linked to the yearly output sheets to produce point prevalence estimates for those years. The estimates for the future years are produced by fitting an epidemic curve (done on the "Set-up" sheet) to the yearly point prevalence estimates.

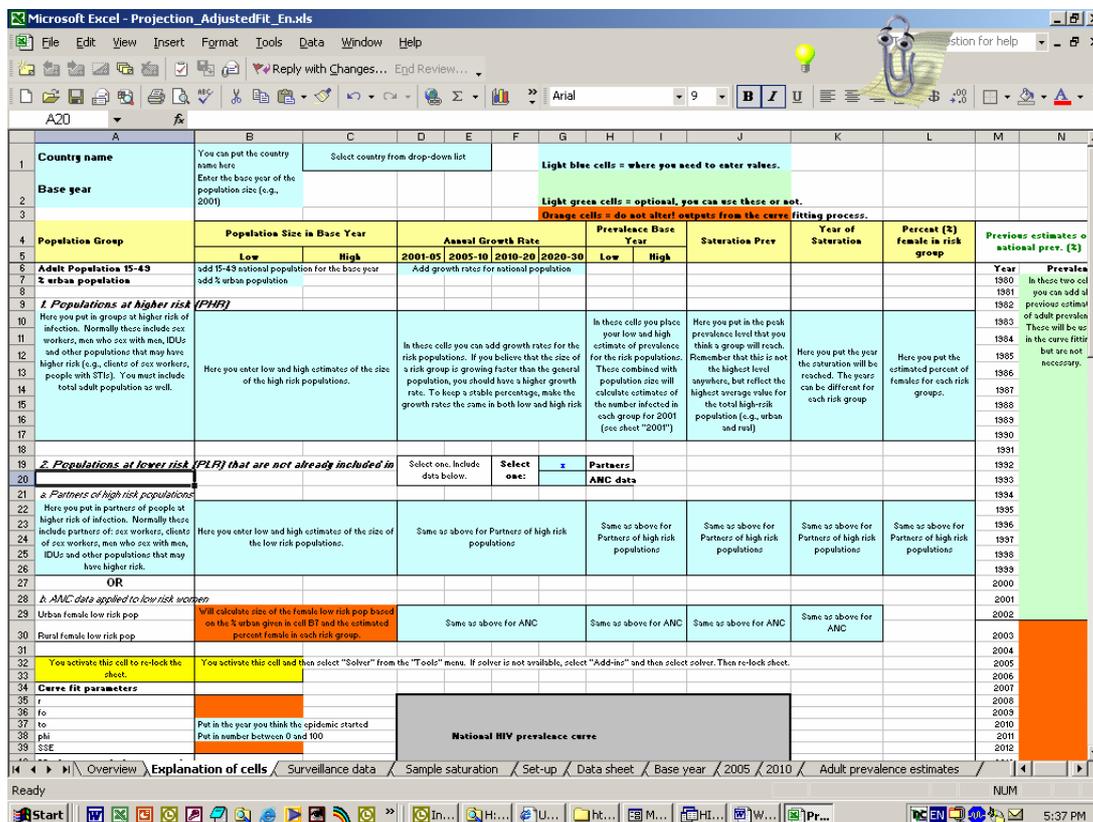
The structure of the Projections Workbook is very similar to that of the Point Prevalence Workbook. You must define populations at higher risk, estimate populations size and prevalence in these populations. However, because you are making projections of prevalence for future years you must also include information about growth, both in terms of the size of the populations and in the prevalence of HIV/AIDS in these groups.



## Overview of sheets in the Projections Workbook

- ◆ The primary sheet is the "Set-up" where all data and assumptions are entered.

- ◆ The sheet titled "Explanation of cells" gives explanations of what cells do and what you should add in the "Set-up" sheet.



- ◆ The sheet titled "Surveillance data" list highest prevalence found in different countries and regions in populations at higher risk. This is background information to help choose appropriate saturation levels.
- ◆ The sheet titled "Sample saturation" shows possible saturation levels and time to saturation for three different types of epidemics.
- ◆ The numbered sheets (Base year, 2005, and 2010) are output sheets that give point prevalence estimates for the year by population.
- ◆ The sheet titled "Adult prevalence estimates" contains the prevalence estimate and projection in a format suitable for copying to Spectrum.

- ◆ The sheet titled "Data sheet" is where you can place the data that were used in the projection as well as the data sources. This information is not automatically linked to the other sheets. It is simply a place to place data and information that you use in making the estimates and projections.

### **C. Basic steps in using the Projection Workbook to develop an epidemic curve and short-term projection of the epidemic**

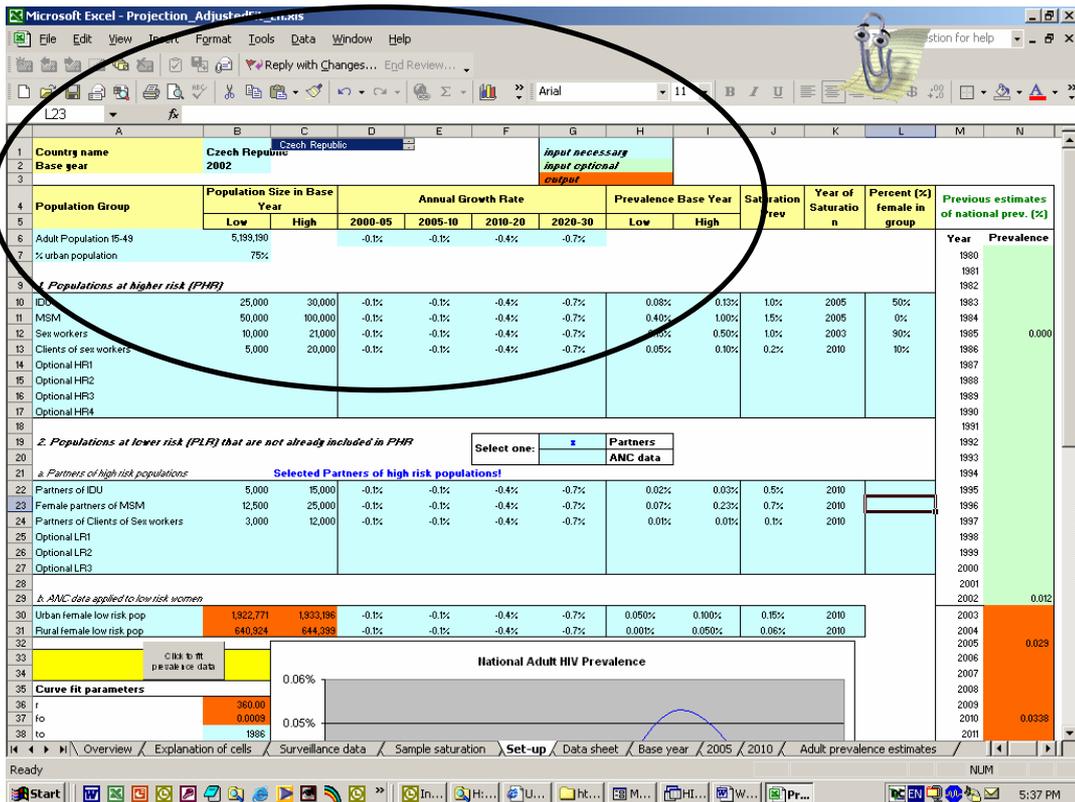
All the information and inputs that will be required to make an epidemic curve and short term projections will be entered in the "set-up" sheet. All the other sheets in the workbook are either explanatory sheets or output sheets. The basic structure of the "Set up" sheet is the same as was used in the Point Prevalence Workbook, but there will be some additional information required to do the short-term projections and to produce the epidemic curve.

#### *1. Base year of estimate and projection.*

This is the current year or the year for which you have the most recent estimate of national prevalence [Cell B2]. You must enter the country or region name [Cell B1] and select it from the list to the right [Cell C1]. This name will automatically appear in the output sheets ("Base Year", "2005", and "2010").

#### *2. Adult population.*

You must also enter the national adult (15-49) population for the base year [Cell B6] and the percent of the population that lives in urban areas [Cell B7]. In addition, you must also put annual population growth for the periods 2000-2005, 2005-2010, 2010-2020 and 2020-2030 [Cells D6 to G6].



### 3. Populations at higher risk.

You must also enter the populations at higher risk [Cells A10 – A17] and estimate the size of the population [Cells B-C10 to B-C17]. These values should come from the most recent year the Point Prevalence Workbook that was used to make a national point prevalence estimate and the base year value (B2) should match those from the Point Prevalence Workbook.

### 4. Population sizes.

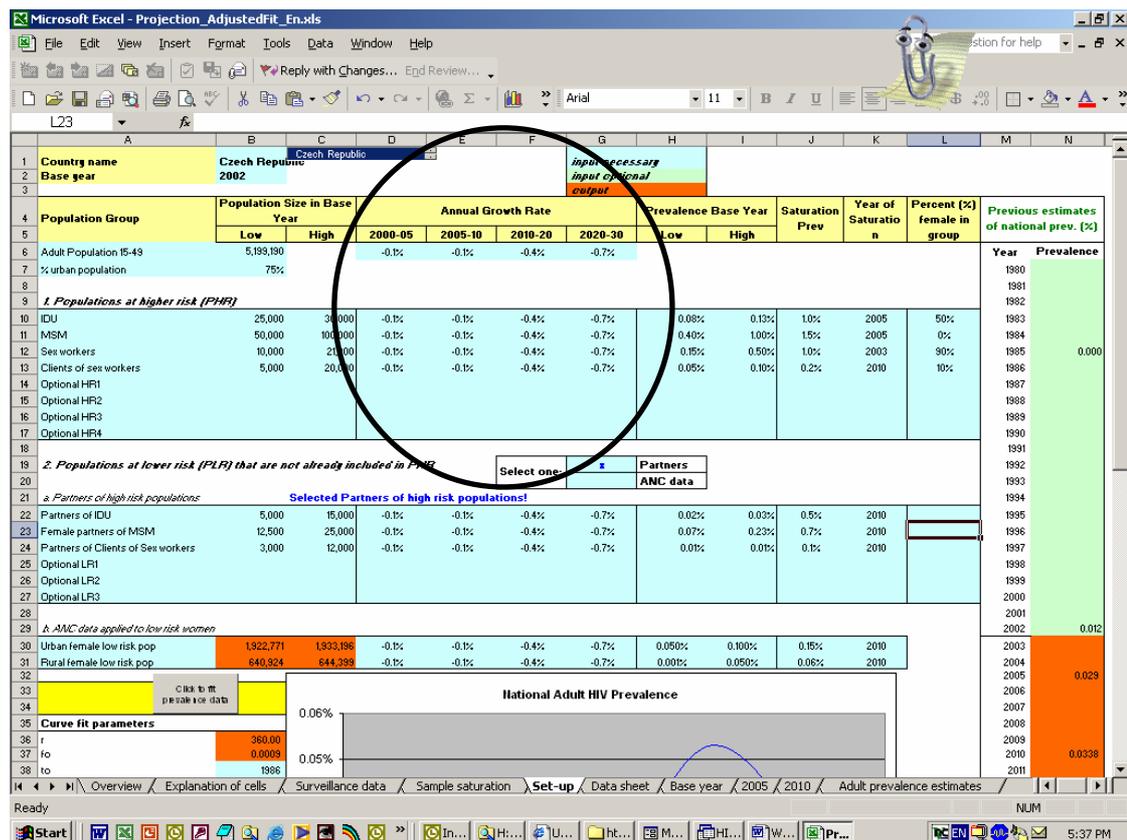
For each of the population groups you must also put in the expected annual growth rate over time [Cells D-G10 to D-E17]. If you expect the size of the populations at higher risk to remain roughly stable as compared to the overall adult populations, you should use the same growth rates as for the overall population [Cells D6 to G6]. If you think these populations are growing more rapidly than the general population, then you can put in those growth values for each population.

5. *Prevalence in populations at higher risk.*

For each of the populations at higher risk you must first enter the estimates of current prevalence among these populations as you did in the Point Prevalence Workbook [Cells H-I10 to H-I17]. For the base year of your projection these values should come directly from the Point Prevalence Workbook.

6. *Growth in prevalence among populations at higher risk.*

For each of the populations you must also set a maximum or saturation HIV prevalence level and the time when that level will be reached. The saturation levels for each of the populations are entered [Cells J10 to J17] as well as the year in which that level will be reached [Cells K10 to K17]. Obviously you cannot know those values with certainty, but you can make educated guesses based on a review of time series data from your country and other countries.

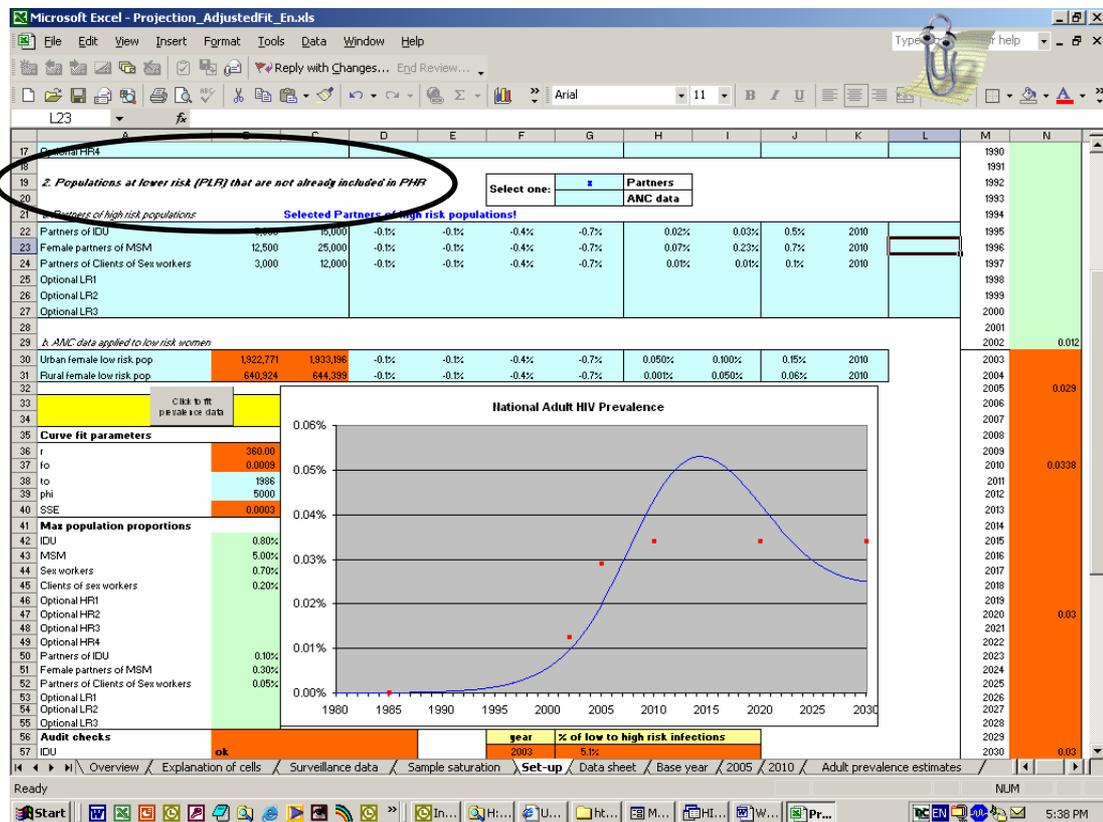


*Note:* There are two sheets in the workbook that contain data that you should review in making the estimates of saturation levels and time, these are “Surveillance data” and “Sample saturation”. The “Surveillance data” sheet has highest HIV prevalence found

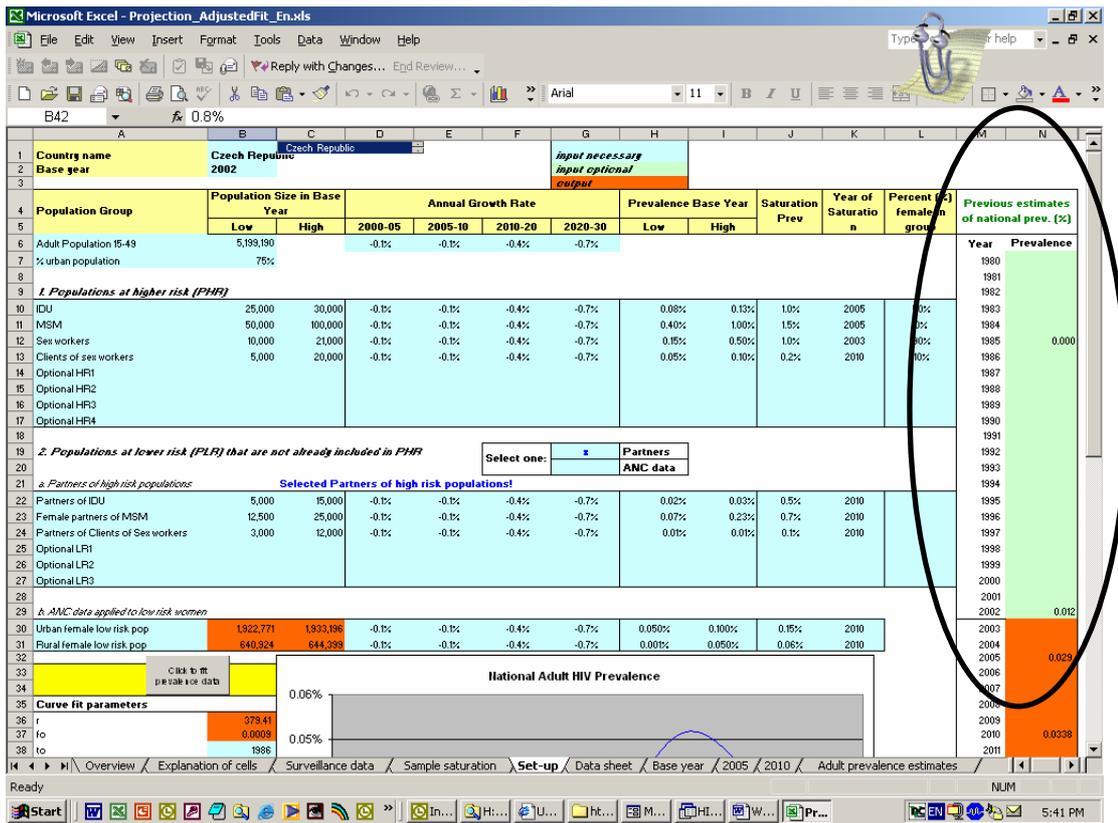
in various populations at highest risk in different countries. These are not meant to be the national saturation level in these populations but they do show what prevalence levels have been reached. The second sheet "Sample saturation" gives some examples of saturation levels and time to saturation that have been used to make projections for different countries. Again, these values are meant to be examples and may not apply to any given country.

7. *Population at lower risk.*

As in the Point Prevalence Workbook, you have two options for handling the prevalence in populations at lower risk of HIV infection. One of the two options must be selected within the "Set-up" sheet. For each option you will need to specify the population size, growth rates over time, base year prevalence, saturation levels and saturation date. As in the Point Prevalence Workbook, you can calculate prevalence in the population at lower risk in two ways, but you must select which way you will use in your final estimate [Cells G19 or G20].



- a) Populations at lower risk (PLR) can be estimated by estimating prevalence in the heterosexual partners of those populations at higher risk. To do this you must estimate their population sizes [Cells B-C 22 to B-C27], expected annual growth rate over time [Cells D-G22 to D-G27], and the low and high prevalence in each group [Cells H-I22 to H-I27], as well as saturation level [J22 to J27], and time to saturation [K22 to K27].
  - b) People with HIV/AIDS from the population at lower risk can also be calculated using HIV prevalence data from pregnant women. Population size for this group is already calculated for women living in urban and rural areas. As with the other populations, you must enter expected growth rate over time [Cells D-G30 to D-G31] prevalence in the base year [Cells H-I30 to H-I31], saturation level [J30 to J31] and time to saturation [Cells K30 to K31].
8. *Entering previous estimates for the curve fitting process.*  
The curve-fitting algorithm will work by fitting a curve to estimates over time as well as projections for future years. The previous estimates must be entered [Cells M7 to M29]. Estimates of prevalence for the base year and projected prevalence for future years will be entered automatically from the information contained in the “Set up” sheet [shown in Cells M30 to M57].



Note: You do not need to have estimates for all years to do a curve fit to develop your epidemic curve. However, you will need at least two previous year's estimates before the base year estimate to produce a reasonable epidemic curve. If you have estimates for more years, include those as well. It is recommended to have national point prevalence estimates from years that reflect both the early and middle stages of epidemic when developing the epidemic curve, not only the most recent years.

9. *Setting the curve fitting parameters, and producing the best-fitting curve.*

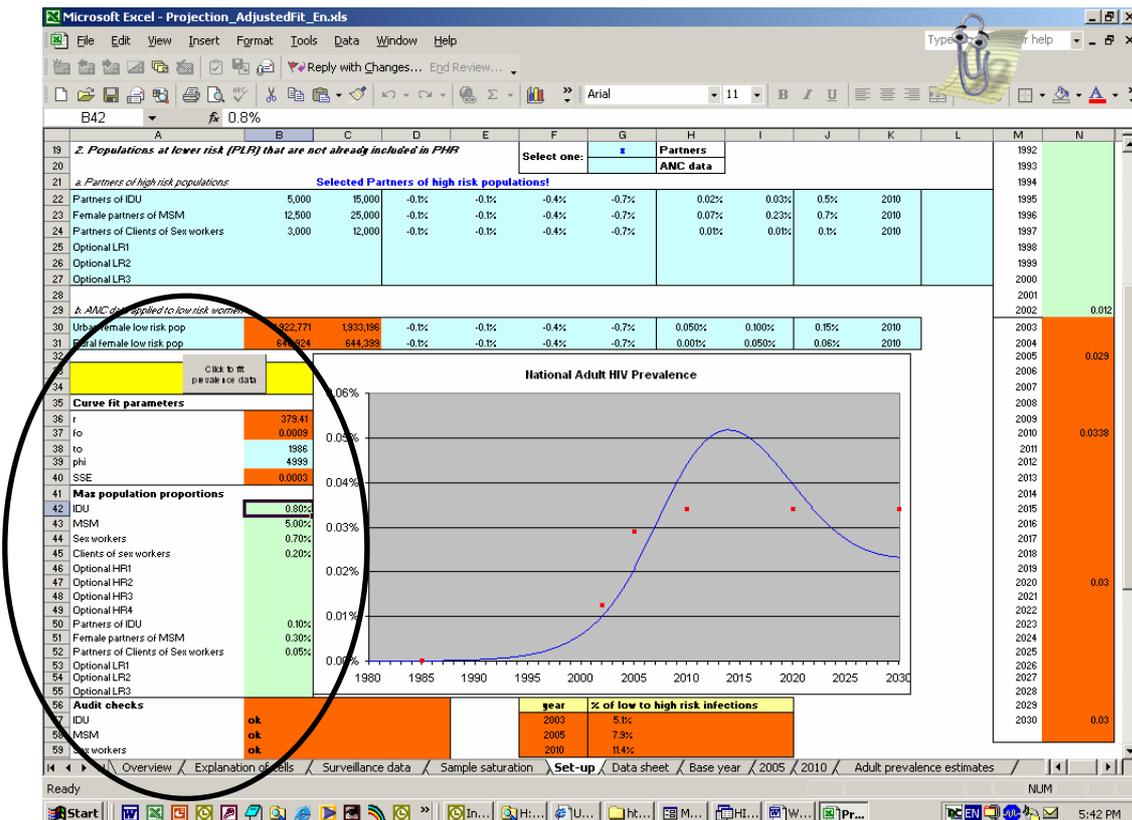
The curve fitting approach uses four parameters to try and produce an epidemic curve for adult prevalence. This approach has been described in detail elsewhere.<sup>1</sup> For the Projection Workbook you must set two parameters in the Curve Fit Parameters sections of the "Set up" sheet [Cells A-B35 to A-B40]. The two parameters you must enter are "t0" [B38] and "phi" [B39]. "t0" refers to the year the epidemic started. "Phi" is a parameter that relates to recruitment of new people into the population. Given that in low-level and concentrated epidemics are made up of populations of

higher risk where people move into and out of the risk populations rather quickly (e.g., sex workers) the standard value to use for “phi” should be set at 1000.

### 10. Running the curve fitter.

Once you have entered the year prevalence estimates and the two default values for the parameters “t0” and “phi” you can then run the curve fitter. This is done by clicking on the button marked “Click to fit prevalence data”. You can see the graphical output of the curve fitter in the section of the spreadsheet marked “National Adult HIV Prevalence”.

*Note:* The curve-fitting algorithm will not always produce on the first try the best fit. If when the curve is produced it does not fit the data points for the yearly estimates, you can re-set the parameters to different values and try re-fitting the curve.



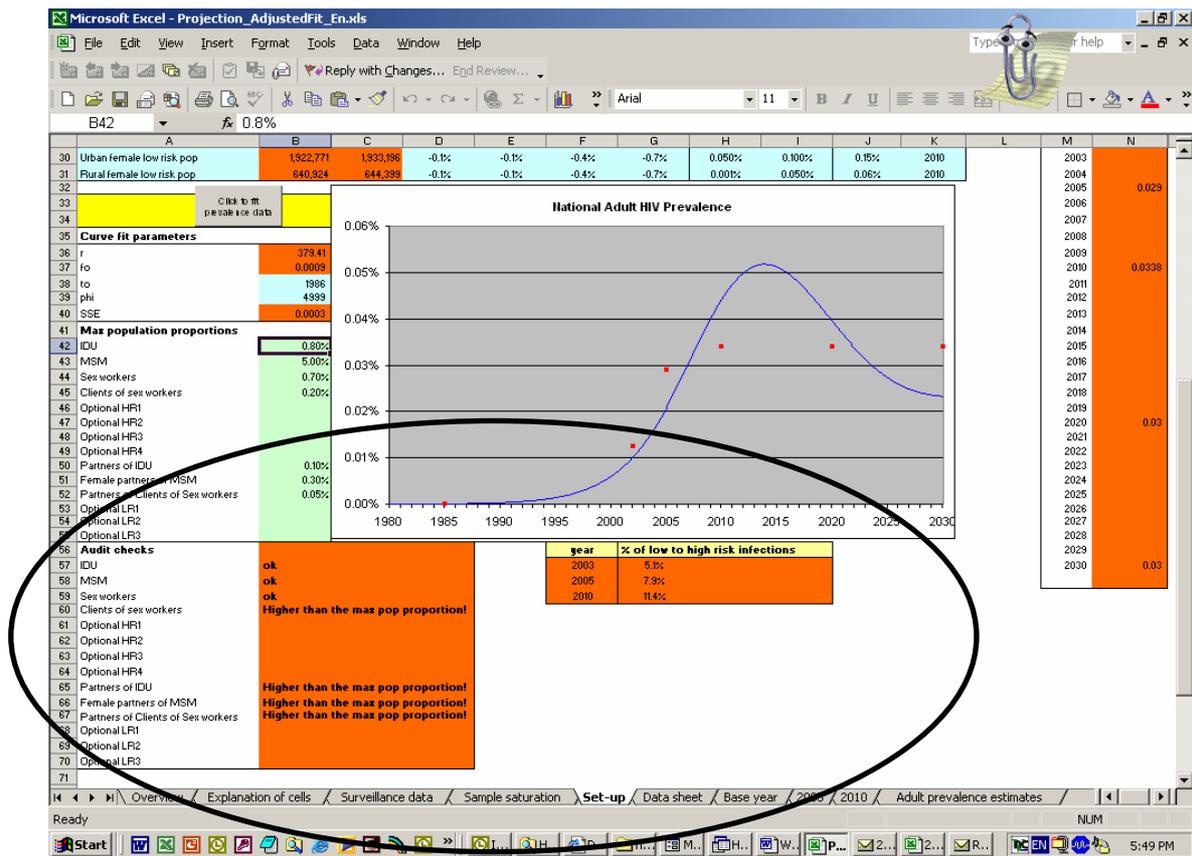
<sup>1</sup> See The UNAIDS Reference Group on Estimate, Modelling and Projections.. Improved methods and assumptions for estimation of the HIV/AIDS epidemic and its impact: Recommendations of the UNAIDS Reference Group on Estimates, Modelling and Projections. *AIDS*, 2002, **16**: W1-W14.

## D. Results and outputs

Once you run the curve fitter, you will have produced an estimate for each year of the epidemic from its start until 2010. The output of this will be in several places. First, in each of the output sheets “Base year”, “2005” and “2010” there will be a spreadsheet that shows the estimates for each population as well as the total population. These have the same structure as that of the Point Prevalence Workbook. There is also a sheet “Adult Prevalence estimates” that gives adult (15-49) HIV/AIDS prevalence over time. This is the output that you will need to import into Spectrum. Each group (low and high) are based on the sums of the number infected in each group divided by the population size of the group.

## E. Checks and reminders after making the initial epidemic curve and short-term projection

The series of consistency checks in the cells A-D56 to A-D70 compare values related to average population sizes of groups at high risk and estimates of prevalence in these groups to ranges found in other countries with low-level and concentrated epidemics. Just as with the consistency checks in the Point Prevalence Workbook these values are not meant to be proscriptive, rather they simply give a comparison to average values. If a value you used in making your national estimate falls outside of this range, a warning is given (e.g., “The percentage of men who have sex with men is unusually low”). If the values are within the range, “OK” is written in the cell.



In addition to these consistency checks there is also a calculation of the percentage of the total estimated number of adults with HIV/AIDS who come from the population at lower risk. This percentage is calculated and displayed for the base year, 2005, and 2010 [Cells F-G 57 to F-G 59 on the “Set-up” sheet]. This can be a key check on the estimate and projection as early in an epidemic almost all of the people estimated to be living with HIV/AIDS should come from the populations at higher risk. Over time that percentage will drop, but in low-level and concentrated epidemics the majority of people living with HIV/AIDS should come from these populations.

Once you are satisfied with the epidemic curve and the short-term projections you can then proceed to develop additional estimates related to HIV/AIDS (e.g., mortality, incidence, HIV among children) using SPECTRUM. This software is described in a separate manual.

**Acknowledgements**

The estimation workbooks and this manual were prepared for a series of regional training workshops conducted in the Spring and Summer of 2003. The workbooks were based on inputs from the UNAIDS Reference Group on Estimates, Projections and Modelling.

Several organizations provided financial and technical support to the development of methods and materials and in the workshops. These organizations were the Centers for Disease Control and Prevention, Family Health International, The East-West Center, The Futures Group, and the Regional Offices of The World Health Organization.