HTC Data Use Tool - User's Manual

Module 1: Inputting data into the HTC Data Use Tool

Table of Contents

How to Use This Module

This module offers a systematic approach to inputting data into the HIV testing and counseling (HTC) Data Use Tool and is meant to accompany the HTC Data Use Tool User's Manual. Although this guide uses only HTC examples, the Tool can be applied and adapted to different data-use questions related to HIV prevention, treatment, and care as well as other programmatic areas.

The manual is designed to provide users with a solid background and understanding of the data synthesis process. It is most useful as a preparation tool for those who will take part in an HTC workshop or those who have completed a workshop and may need to refer back to the concepts and processes at a later time.

All materials related to this manual and the HTC Data Use and Strategic Planning Tool can be found here: http://globalhealthsciences.ucsf.edu/prevention-public-health-group/global-strategic-information-gsi/monitoring-and-evaluation/hiv-testing

1. Introduction to inputting data into the HTC Data Use Tool

The HIV Testing and Counseling (HTC) tool is designed to answer key strategic questions related to HTC services through enhanced visualization of data. Utilizing Excel and GIS software, the tool can organize data, display tables and graphs and produce maps that aid strategic planning and policy decisions.

Data used with the HTC Data Use Tool are input into the *2-Inputs* tab of this Excel-based tool. In this tab, columns (display data vertically) represent indicator fields, while rows (display data horizontally) represent single records. The first row of this tab is a header that contains the column labels as shown in Figure 1.

As described in this module, data is input into the tool by completing the following steps:

1. Creating a Data Entry Plan
2. Double data entry
3. Checking and rectifying double data entry
4. Creating a Data Key

Figure 1. Example of *2-Inputs* tab in HTC Data Use Tool

2. The Data Entry Plan

## 2.1.	Inventory and gather all available data

The Data Entry Plan is an Excel spreadsheet that contains information on all available data and will guide the data entry process. The first step in creating a Data Entry Plan is to inventory and gather all available data. Data to be input into the tool should be of various sources (e.g. Population based surveys, Surveillance data , Program Data, Census data, etc.), time points (e.g. most recent year + past years to show trends) and at a minimum the following stratifications: Geography (National, Province, District); Sex (male, female, both) ; Age group (<15 and 15+ at a minimum). Data should be entered for the most *disaggregated level* (e.g. if data at the district level is available, enter district data). See Appendix 1 for the *HTC Suggested Indicator* list.

## 2.2.	Developing a Data Entry Plan

This Data Entry Plan should contain the following Columns:
- Column A: Source and year
- Column B: Indicator
- Column C: Geographic disaggregations (National, Province, or District)
- Column D: Gender dissagregations (male, female, or both)
- Column E: Age disaggregations
- Column F: Source location (if published source list the page and table number for future reference)

As a group, decide which indicators to include in the tool from each source and complete the Data Entry Plan accordingly. Each row should have text in all columns so that each disaggregation appears as its own row in the spreadsheet even if duplicated as shown in Figure 2 below.
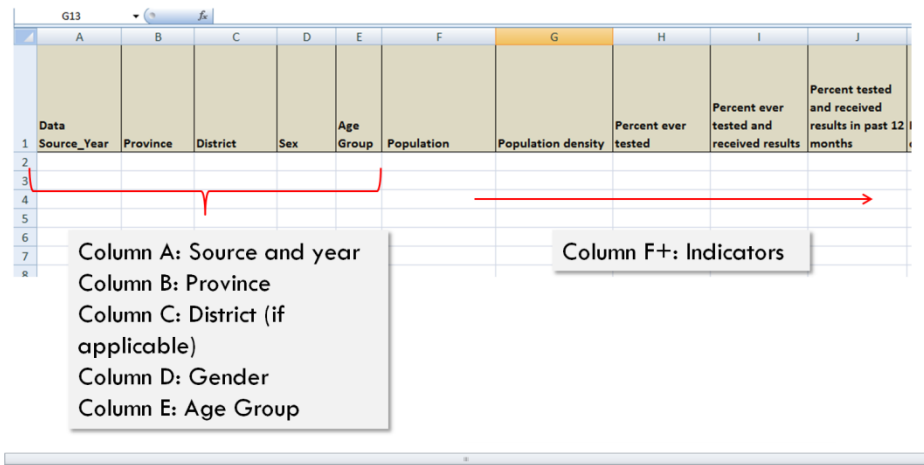
Figure 2.

	A	B	C	D	E	F	G	H
1	Source	Indicator	Geographic disagreagations	Gender disaggregations	Age diaggregations	Source location (if published source)		
2	HMIS 2012	Antenatal 1st visit before 20 weeks	National	All	< 20weeks	HTC indicators and Data set		
3	HMIS 2012	Antenatal 1st visit before 20 weeks	Province	All	< 20weeks	HTC indicators and Data set		
4	HMIS 2012	Antenatal 1st visit before 20 weeks	District	All	< 20weeks	HTC indicators and Data set		
5	HMIS 2012	Antenatal 1st visit 20 weeks or later	National	All	20 weeks or later	HTC indicators and Data set		
6	HMIS 2012	Antenatal 1st visit 20 weeks or later	Province	All	20 weeks or later	HTC indicators and Data set		
7	HMIS 2012	Antenatal 1st visit 20 weeks or later	District	All	20 weeks or later	HTC indicators and Data set		
8	HMIS 2012	Antenatal 1st visit by woman <18 years	National	Female	< 18 years	HTC indicators and Data set		
9	HMIS 2012	Antenatal 1st visit by woman <18 years	Province	Female	< 18 years	HTC indicators and Data set		
10	HMIS 2012	Antenatal 1st visit by woman <18 years	District	Female	< 18 years	HTC indicators and Data set		
11	HMIS 2012	Attendance child health <1 year female	National	Female	<1 year	HTC indicators and Data set		
12	HMIS 2012	Attendance child health <1 year female	Province	Female	<1 year	HTC indicators and Data set		
13	HMIS 2012	Attendance child health <1 year female	District	Female	<1 year	HTC indicators and Data set		
14	HMIS 2012	Attendance child health <1 year male	Province	Male	<1 year	HTC indicators and Data set		
15	HMIS 2012	Attendance child health <1 year male	District	Male	<1 year	HTC indicators and Data set		
16	HMIS 2012	Attendance child health <1 year male	National	Male	<1 year	HTC indicators and Data set		
17	HMIS 2012	Attendance child health 1-4 years female	National	Female	1-4 years	HTC indicators and Data set		
18	HMIS 2012	Attendance child health 1-4 years female	Province	Female	1-4 years	HTC indicators and Data set		
19	HMIS 2012	Attendance child health 1-4 years female	District	Female	1-4 years	HTC indicators and Data set		
20	HMIS 2011	Antenatal 1st visit before 20 weeks	National	All	< 20weeks	HTC indicators and Data set_2010_2011		
21	HMIS 2011	Antenatal 1st visit before 20 weeks	Province	All	< 20weeks	HTC indicators and Data set_2010_2012		
22	HMIS 2011	Antenatal 1st visit before 20 weeks	District	All	< 20weeks	HTC indicators and Data set_2010_2013		
23	HMIS 2011	Antenatal 1st visit 20 weeks or later	National	All	20 weeks or later	HTC indicators and Data set_2010_2015		

data entry plan 2014 / Sheet2 / Sheet3

3. Double data entry

Once all data has been accounted for in the Data Entry Plan, two different individuals will transfer data from its original source into two separate spread sheets that resemble the format of the HTC Data Use Tool. To start, create one HTC Data Use Tool skeleton matching the spreadsheet shown in Figure 3 below. Customize the geographic stratifications as appropriate.

Figure 3.

Next, make a copy of this skeleton and paste into a new workbook by right clicking on the 'Sheet 1' tab and following the steps shown in Figure 4 below. Save one spreadsheet as 'Tool 1' for one person to complete, and the other spreadsheet as 'Tool 2' for the other person.

Figure 4.
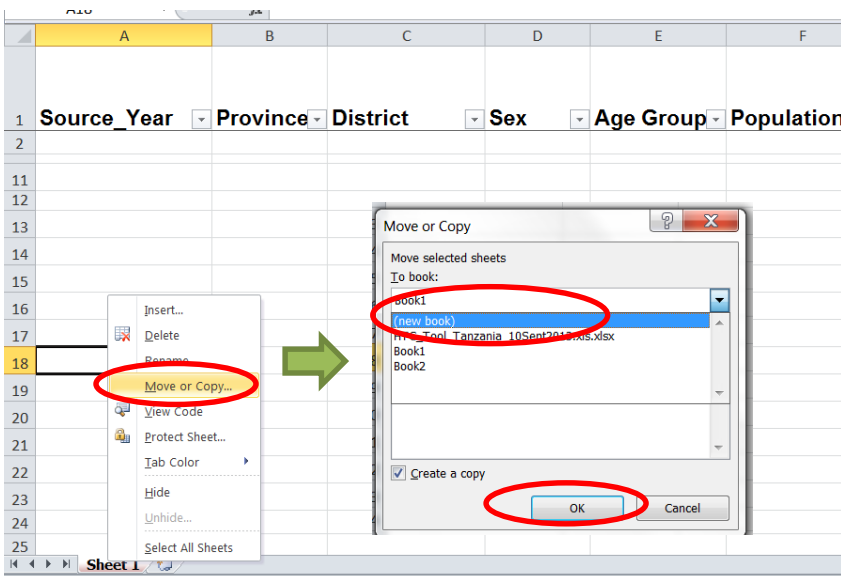
Each data entry representative should enter data into their respective tools. Disaggregations (e.g. Columns A-E: sex, age group, etc.) and indicator text should be entered exactly as they appear in the Data Entry Plan. Columns A-E should have text in *every row* of the tool with data whereas columns F+ will only have data for that indicator (column) as it is available for that disaggregation (row). The data in Tool 1 will be compared to the data in Tool 2 in a later step so it is important that all text (e.g. numbers, spelling, etc.) in both tools are exactly the same.

Data entry tips:
- For National-level indicators, enter 'National' into the Region/Province column
- For Province-level indicators, leave the District column blank
- Convert all PDF tables to Excel spreadsheets so tables can be copied and pasted instead of typed from the source to the skeleton

After all data has been entered, both Tool 1 and Tool 2 should resemble Figure 5 below.

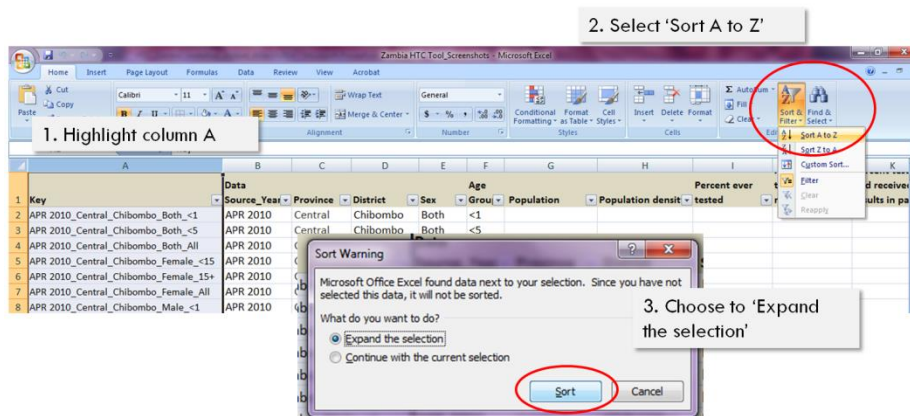**Figure 5.**

4. Checking and rectifying double data entry

After both data entry representatives have entered all data into their respective Tools, create a 'key' from the contents of columns A-E to compare the two tools. In both Tool 1 and Tool 2, insert a new Column A and in cell A2, enter the following formula: =B2&"_"C2&"_"D2&"_" E2&"_"F2 which condenses the contents of these data disaggregation cells into one cell. Copy this formula by dragging it from cell A2 down to the last row of the Tool that contains data as shown in Figure 6. The key will be used to sort the data so data in each tool is in the same order.

Figure 6.
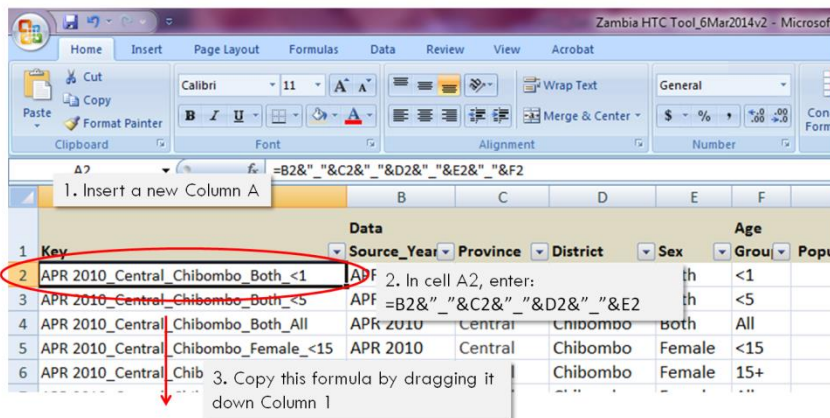
Next, sort both Tool 1 and Tool 2 by the key (column A) by highlighting column A, selecting 'sort A to Z' from the Home tab and choosing to 'Expand the selection' as shown in Figure 7. Now the contents of Tool 1 are displayed in the exact same order as the contents of Tool 2.
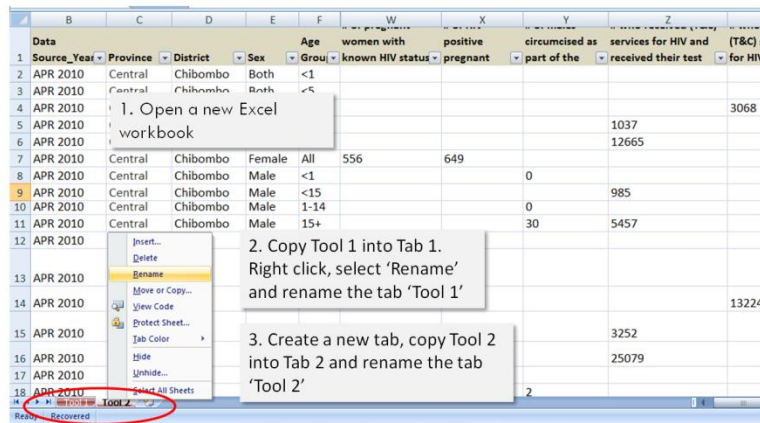
Figure 7.

To compare the contents of the two Tools, first open a new workbook and copy Tool 1 and Tool 2 into separate tabs in this new workbook. Rename each tab 'Tool 1' and 'Tool 2', respectively, as shown in Figure 8.

Figure 8.

Next, use formulas to identify any data entry discrepancies between Tool 1 and Tool 2. To do this, first create a new tab and rename it 'Check'. In cell A1 of the 'Check' tab enter the following formula: ='Tool 1'!A1='Tool 2'!A1. Copy this formula across all cells that have contents in Tool 1 and Tool 2. For example, if Columns A-AG and rows 1-500 have contents then drag this formula down to row 500 and across to column AG as shown in Figure 9. All cells that match exactly in Tool 1 and Tool 2 will appear as 'TRUE' in the 'Check' tab; any that do not match will appear as 'FALSE.
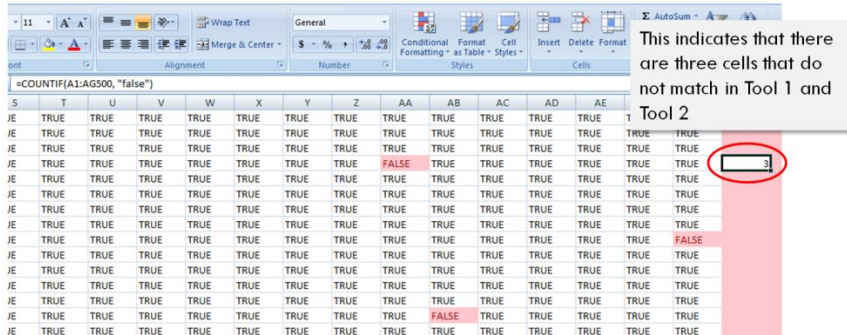
Figure 9.
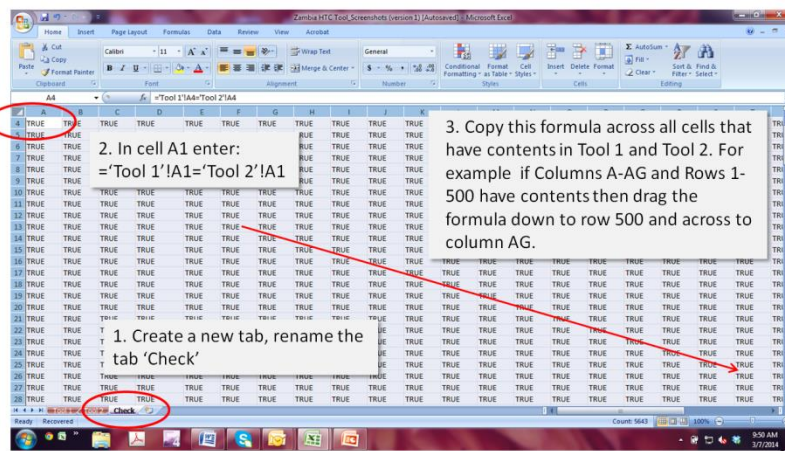
To help identify which cells do not match between the two tool skeletons, highlight all 'FALSE' cells in red. To do this, first select all cells in the spreadsheet, then select under the 'Home' tab Conditional Formatting => Highlight Cell Rules => Equal To. In the 'Equal To' dialog box type 'false' to specify that you want to Format cells that are equal to 'false' with Light Red fill and Dark Red Text as shown in Figure 10.

Figure 10.
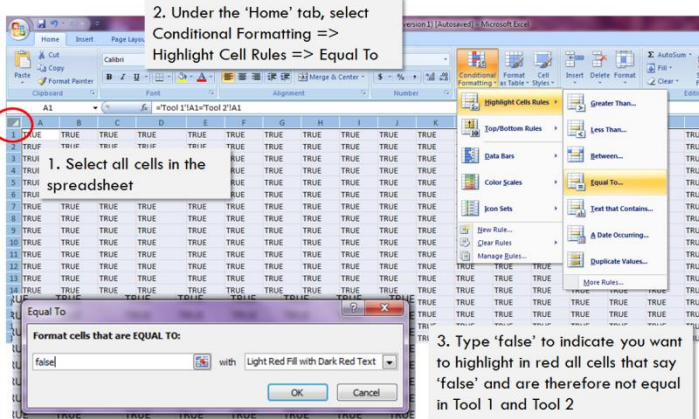
To identify how many cells do not match between Tool 1 and Tool 2, set a count by selecting any empty cell in the 'Check' spreadsheet and entering the formula: =countif(A1:AG500, "false") where 'AG500' is the cell furthest down and furthest to the right with contents in Tool 1 and Tool 2 as shown in Figure 11. The number that appears in this cell indicates the number of cells that do not match between the two tools. As errors are corrected, this number will reduce to 0, indicating no data discrepancies exist between Tool 1 and Tool 2.

Figure 11.

To identify and correct errors between the two tools, first select a cell containing 'FALSE'. The formula bar shows which cell in each of the tools is discrepant. As shown in Figure 12 below, the contents of cell AA4 in Tool 1 and Tool 2 do not match. From here, look back at the original data to determine what the correct value for that cell is and update it in the Tool 1 and/or Tool 2 tabs. Once the contents of this cell match in Tool 1 and Tool 2, it will change to 'TRUE' and the count will be reduced by one in the 'Check' tab.

Figure 12.
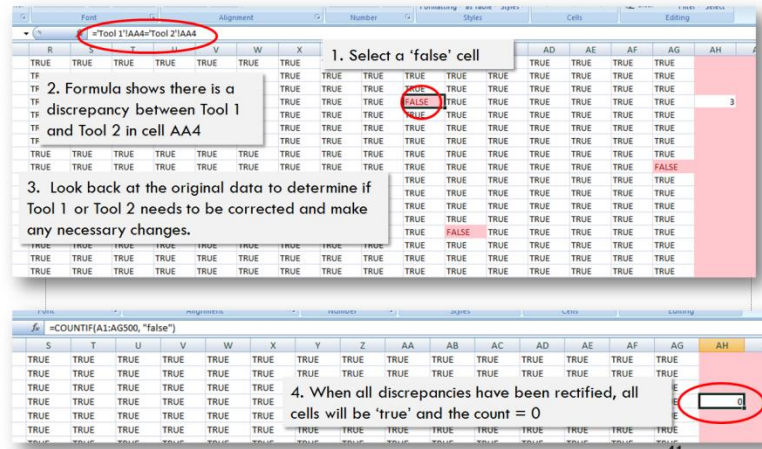
After all errors have been rectified it can be assumed that Tool 1 and Tool 2 contain accurate data that exactly reflect the source data. From here you can save the contents of Tool 1 into the Master HTC Data Use Tool. To do this, highlight and copy the contents of Tool 1, excluding column A (the key). In the Master HTC Data Tool, select the '2-inputs' tab and paste these data as shown in Figure 13.
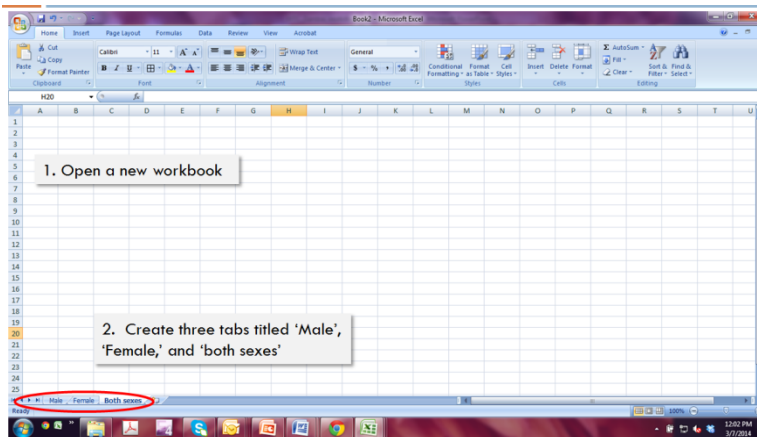
Figure 13.
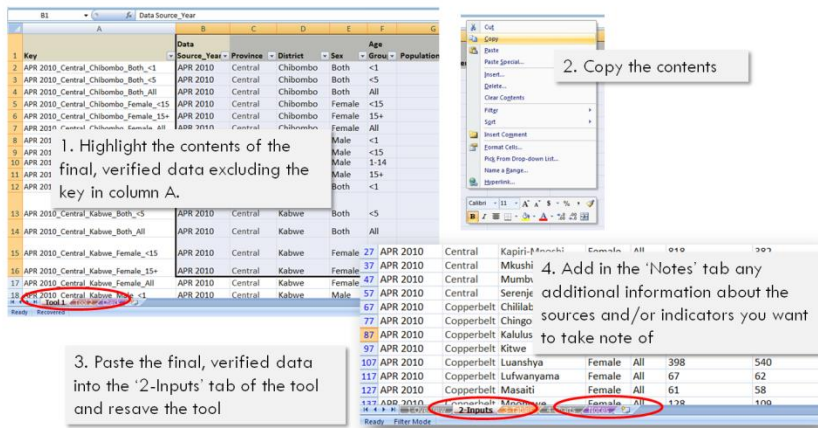
5. The Data Key

The Data Key is an Excel-based spreadsheet guide that shows which indicators are included in the HTC Data Tool, including their respective disaggregations and sources. This dynamic guide will be used as a reference when creating graphs and maps from this data. As such, it should be revised when updating the HTC Data Use Tool.

The Data Key is separated by three tabs, each titled by the sex disaggregations in the tool. Within each tab, Column A lists the indicators that contain data for that sex, row 1 describes age-disaggregations, while row 2 describes geographic disaggregations. The contents of this table indicate the source from which these data come.
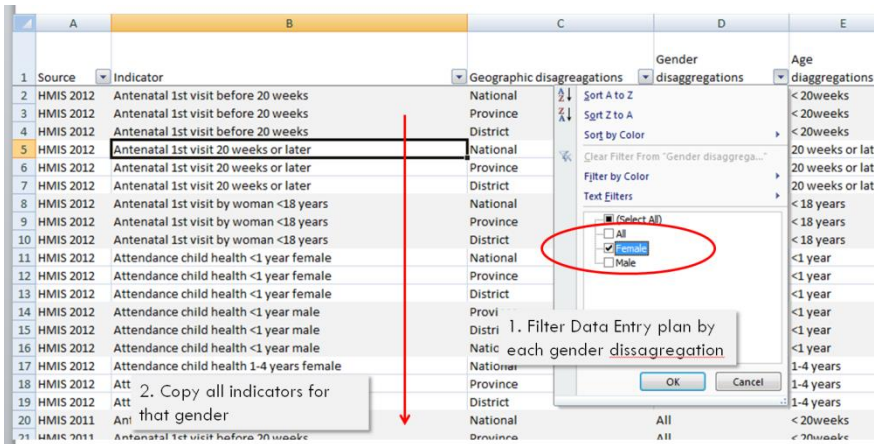
To start, open a new Excel workbook and create three tabs renamed 'male', 'female', and 'both Sexes' as shown in Figure 14.

Figure 14.

Returning to the Data Entry Plan, filter the data by each sex and copy all indicators with data in the Tool by that sex as shown in Figure 15.

Figure 15.

Paste this list of indicators into Column A of the respective tab in the Data Key workbook (e.g. indicators filtered by female are pasted in the 'female' tab). Delete all duplicate indicators/rows as shown in Figure 16.

Figure 16.
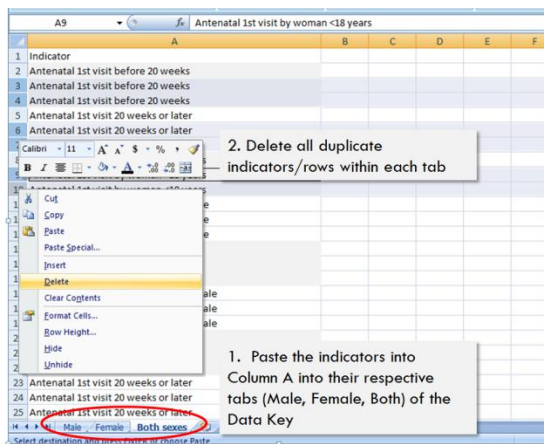
Lastly, use the filter function in the Data Entry Plan to fill in the remaining contents of each tab. For example, filter by various sexes, ages and geographic disaggregations to determine which indicators are represented and by which source. This is an iterative process and will take some time to develop an effective, customized strategy. Figure 17 is an example of a completed Data Key.

Figure 17.

Appendix 1. HTC Suggested Indicator List

Indicators to be collected prior to use of the tool include:

Data Source	Indicator	Stratifications
Population based surveys (DHS, AIS etc.)	- HIV prevalence - % and # ever HIV tested - % and # HIV tested in the past 12 months - % and # HIV tested in the past 12 months who received their results - % and # received HIV testing and counseling in the past 12 months - % and # HIV tested, counseling and received results in the past 12 months	- Geography (National, Province, District) - Sex (male, female, both) - Age group (<15 and 15+ at a minimum)
Surveillance data (ANC, PMTCT, BSS etc.)	- HIV prevalence - # HIV positive - % and # attending ANC with unknown HIV status at entry	- Geography (National, Province, District, Sub-district, Facility) - Sex (male, female, both) - Age group (<15 and 15+ at a minimum)
Program Data (APR, HTC etc.)	- # HIV counseled and tested - # HIV counseled, tested and received results - % and # couples counseled and received results - % and # HIV-positive pregnant women who received antiretrovirals for PMTCT - ART treatment coverage (estimated # of persons eligible for treatment / estimated # of persons receiving treatment)	- Geography (National, Province, District, Sub-district, Facility) - Sex (male, female, both) - Age group (<15 and 15+ at a minimum) - HTC service type (e.g. PMTCT, TB, mobile, outreach, etc.)
Census data	- Population projections - # HIV positive	- Geography (National, Province, District, Sub-district) - Sex (male, female, both) - Age group (<15 and 15+ at a minimum)
Geographic data	- Shapefile for administrative levels of interest	

Recommended indicators:

Data Source	Indicator
Surveillance data	- HIV incidence - Sexually transmitted disease prevelance - HIV prevalence among persons with STI symptoms in the past 12 months - TB/HIV co-infection prevalence - TB prevalence
Program Data	- Linkages to care and treatment (# newly diagnosed as HIV+, # newly enrolled in care) - VCT essential services coverage - PMTCT essential services coverage - Availability of HIV/AIDS services by facility (HIV testing, Any care and support, any PMTCT, ART, PEP, Youth friendly VCT/PMTCT) - Coverage of PMTCT (# annual births, # pregnant women with 1+ ANC visit, # pregnant women receiving PMTCT including testing, # receiving complete course of ART)
Behavioral data	- Awareness of HIV-status - Perceived risk of HIV infection - Prevalence of HIV risk-related behaviors (# of sexual partners, condom use, circumcision, age at first sex)
Geographic data	- Geographic size of regions (sq km) - Geo-coordinates for HTC health facility sites (e.g. ART, HTC site locations etc) - Road shapefiles