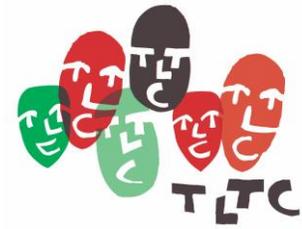


# DOCUMENTATION GUIDE #15: TLT-2



Prepared by Stephanie Chamberlin

*Latest Version: September 12, 2019*

## **TLT-2 OVERVIEW**

TLT-2 survey data were collected in 2015, three years after the last round of TLT-1 data collection (see Cohort Profile). This follow-up survey was designed to assess changes in HIV and fertility relationships in the rapidly evolving HIV testing and treatment landscape in Malawi between 2012 and 2015.

### **Sample**

80 percent of women (N=1200) and 71 percent of men (N=407) from the 'core sample' of TLT-1 were re-interviewed in 2015 (see Cohort Profile). An additional 253 women from the refresher sample that immediately followed Wave 8 were also re-interviewed in 2015 (also a 80% response rate), for a total of 1453 women in the TLT-2 dataset. Again in this phase of the study, women were given tokens to share with their sexual and romantic partners, allowing these men to enroll in the study. Male partners from TLT-1 were only included in TLT-2 if they were still in a relationship with a female respondent (N=311). Women recruited 262 new, current male partners using tokens (total N=573) (see Couples Documentation Guide).

Women and male partners who appear in both the TLT-1 and TLT-2 datasets can be followed longitudinally across study phases using a single, unique identifier. Respondent identifiers ('respid' in the dataset) are coded to indicate to which part of the sample the respondent belongs (see Documentation Guide #3 on Respondent IDs and Missing Codes and Couples Documentation Guide #13). Random men who were enrolled as part of the core sample in 2009 have respids that begin with "5". Male partners enrolled during waves 1-8 of TLT-1 were assigned respids that begin with "6". Male partners who enrolled for the first time in TLT-2 have respids that begins with "7".

### **Variable Consistency**

This follow-up survey included many of the same variables captured in TLT-1. However, some variables were adapted slightly, new questions and sections were added, and some variables or sections of the questionnaire were dropped. We encourage analysts to review the 'Data Key' to help track continuity in survey items over time. This document can be found on the TLT study website. When question wording varies by gender, the variables are given different suffixes (e.g., lm6 for women vs. lm6m for men in TLT-2). Likewise, where similar questions are asked differently, we used suffixes to distinguish the measures from one another and prevent accidental merging (e.g., land ownership for the household as e22 vs personal land ownership as e22b). In other instances, the content or exact question from TLT-1 has been moved to a different section of the questionnaire in TLT-2, and therefore, the variable name is different while its content remains the same (e.g., b1 in TLT-1 is the same as y5 in TLT-2). Again, the 'Data Key' can help users identify these cases. Although we have tried to be thorough and consistent with naming commensurate variables and differentiating differences using suffixes, analysts are responsible for checking the questionnaires for consistency rather than assuming harmonization.

## Mode of Survey Implementation

Both TLT-1 and TLT-2 data were collected by trained interviewers, in-person. In 2015, most interviews were conducted in small private rooms at the TLT research center in town. In a small number of cases (N=32), interviews were conducted via telephone because the respondent had moved or was working away from Balaka. For other migrants for whom we did not have a working telephone number, a migration team was charged with finding them by asking family members and friends where they moved. When possible, these migrants were followed up in person and interviewed in their new homes or another convenient location. The location of the interview is captured in the `i7` variable. When migrants could not be found, interviewers asked a family member or friend to answer a migration autopsy, a short survey that describes when and where the person moved (see separate migration autopsy dataset).

In TLT-1, survey data were captured on a paper questionnaire and later entered into a database by data entry staff. During TLT-2, the questionnaire was programmed into electronic tablets using ODK (Open Data Kit) software. All survey instructions, questions, and responses appeared on the tablet; the order of the survey and all skip patterns were enforced by the tablet itself. This process helped to decrease data entry time, limited data entry errors, and decreased interviewer confusion about survey instructions or skip patterns.

The published paper version of the questionnaire reflects the order and instructions that were used to administer the electronic version of the survey tool.

## Early Data Collection Issues

During the first two weeks of data collection (July 1- July 15, 2015) for TLT-2, a few glitches with the electronic survey were identified and fixed. A few of these glitches resulted in missing data on the men's survey. If analysts note any data anomalies, such as a cluster of missing values, it may be useful to confirm that these issues were associated with the timing of survey implementation using the 'interviewdates' (string) or 'interviewdaten' (numeric) variables. Please report any anomalies not covered here on the TLT data [user forum](#).

## SPECIFIC ISSUES

### Section M: Missing marriage histories

Due to the combination of interviewer error and a problem with the survey programming on tablets 124 men and 2 women are missing marriage histories. A programming mistake allowed the question `m4_1` (partner name, which is now redacted from the public datasets) to be left unanswered, triggering a skip ahead to `m9a`. The error affected questions `m4-m8` but not `m1-m3` or `m9a-e`. For women and for men that took part in TLT-1, it may be possible to recreate much of the marriage history using TLT-1 data and current marital status from TLT-2. As shown below, the problem is clustered by interviewer.

Random Men:

```
tab interviewer if m5_1==.m & m1~=.5, nol
```

Interviewer	Freq.	Percent	Cum.
202	1	4.76	4.76
204	1	4.76	9.52
205	17	80.95	90.48
206	1	4.76	95.24
208	1	4.76	100.00
Total	21	100.00	

Partnered Men:

```
tab interviewer if m5_1==.m & m1~=5, nol
```

Interviewer	Freq.	Percent	Cum.
203	1	0.97	0.97
205	70	67.96	68.93
206	28	27.18	96.12
207	2	1.94	98.06
208	2	1.94	100.00
Total	103	100.00	

Women:

```
tab interviewer if m5_1==.m & m1~=5, nol
```

Interviewer	Freq.	Percent	Cum.
209	1	50.00	50.00
210	1	50.00	100.00
Total	2	100.00	

### Section CS: Card Sort missing data

Data on the proper allocation of anti-retroviral medicine were collected using visual prompts in the form of picture cards. A small number of interviews were conducted via phone (see mode of survey section above) rather than face-to-face interviews. The Card Sort Section was skipped for those respondents who were interviewed by phone because they were not able to see the visual prompts. 13 women and 19 men are missing data on the Card Sort Section, all of whom responded to survey via telephone. The location of the interview is captured in the *i7* variable.

### Section Y: Y2 and Y15 Issues with travel time

Analysts should use the variables *y2*(a and b) and *y15* with caution. These variables both aim to capture the length of travel time to the interview location (*y2*), and to the nearest trading center (*y15*). The main interview site is located near the main training center in Balaka. The research team has noted a number of inconsistencies and implausible responses to these two questions. First, there are units of time written as decimals in *y15*, such as 1.2 and 2.3, which is an uncommon way to discuss time, especially in units of minutes. While the question asked for travel time in minutes, it is likely that these responses were referring to hours. It is challenging to identify if other responses given as whole numbers may also have been referring to hours instead of minutes. Cross-checking *y2* and *y15* may provide some insight, as *y2* allowed responses in minutes or hours, although this does not fully solve the issue. Second, there are both high and low outliers that indicate improbable travel times based on local knowledge of the respondent's listed village. Similar issues were identified for variable *tr3* in TLT-1. These inconsistencies have been left in the data; we leave it to analysts to make their own judgement calls about how to assemble these data and correct the anomalous values since there is no single correct means of imputing these values. Shared solutions may be posted on the user forum.

### **Section CR: Child Roster pre-population differences between men and women**

While randomly-selected women were followed from TLT-1 to TLT-2, new male partners were added at TLT-2 and non-current male partners from TLT-1 were not included in TLT-2 (see the Couples Documentation Guide). For the Child Roster Section, this meant that data on children from TLT-1 could be pre-populated into the survey tablet for women, but not for men.

For children born before 2012 and reported during TLT-1, only those still living at a woman's last interview were pre-loaded for interviewers to verify. Women were asked to report on new information for these children and any additional children born in or after 2012 or not reported during TLT-1 (if, for example, a woman missed many waves of TLT-1). Men had no information pre-populated into the survey tablet from TLT-1. All men were asked to report on all children born alive, even if they were already reported during TLT-1. Thus, children who had died prior to 2012 were excluded in the women's data but are included in the men's 2015 reports. A more comprehensive dataset of all births (women and men) is available as part of a separate TLT Births dataset.

### **Section LM: Skip logic inconsistencies in women's data**

For women, interviewers were instructed to ask the LM section (LMUP) of respondents who had been pregnant or given birth since 2013. This skip logic was not enforced by the tablet, leading to a few errors. 24 women who should have been asked these questions were accidentally skipped. These women are coded as missing (.m) on the LM questions. In contrast, the 663 women who did not report a pregnancy or birth since 2013 are coded as legitimate 'skips' (.s). Men were specifically asked the date of the last pregnancy in the lm0 variables, and the skip logic was automatically applied in the tablet based on lm0y.

### **Section Q: Health Services**

The format of collecting data on health services changed between TLT-1 and TLT-2. Analysts using the TLT-1 data in combination with TLT-2 data should proceed with caution and read the associated questionnaires and documentation guide (#12 Health Services) carefully.

---

<sup>i</sup> Yeatman, S., Chilungo, A., Lungu, S., Namadingo, H., & Trinitapoli, J. (2019). Tsogolo la Thanzi: A Longitudinal Study of Young Adults Living in Malawi's HIV Epidemic. *Studies in family planning*, 50(1), 71-84.