

2020

Add Health Wave IV Documentation



**Report prepared by**

Ping Chen

Kathleen Mullan Harris

# Wave IV Constructed Current Relationship Status



CAROLINA POPULATION CENTER | CAROLINA SQUARE - SUITE 210 | 123 WEST FRANKLIN STREET | CHAPEL HILL, NC 27516

Add Health is supported by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations.

## Acknowledgment

### When using these data, please cite:

Ping Chen and Kathleen Mullan Harris. 2019. Wave IV Constructed Current Relationship Status. Carolina Population Center, University of North Carolina at Chapel Hill.

## Introduction

Add Health Wave IV data provide detailed information about the romantic relationship history of each respondent. The relationship sections, including Section 16A, 16B, and 17, have a complex data structure. One needs to incorporate a list of variables across sections and go through multiple steps of data management to construct a current relationship status variable. In order to facilitate the use of relationship data, we decided to release the Wave IV current relationship status variables Chen and Harris developed for their research projects.

This documentation describes how Wave IV current relationship status and related flag variables were constructed.

## Data Description and Variable Construction Steps

We used Wave IV data from three sections to draw information, including Section 16A, 16B, and 17. Section 16A is a summary of relationship data and covers Questions 1 to 12. Each respondent (AID) has one observation in the data file. Section 16B is a stacked data file and arranged by AID and partner ID (PRTN\_ID) and covers Questions 13 to 30. This means each respondent can have multiple partners and each partner has relationship data attached to it. Section 17 is about one relationship per respondent (details will be provided in the later part of the section). Our ultimate goal was to select information from these three sections to define the current relationship status for each respondent.

Here we lay out the data management steps that fulfilled our goal.

### Step 1.

We used Section 16A, a summary of relationship data file per respondent, to create the first flag variable (FLAG1). FLAG1 is able to define respondents who have never had a romantic relationship by Wave IV.

### Step 2.

We used Section 16B, a stacked relationship data file (meaning that each respondent may have multiple partners), to find out the respondent's relationship status with each partner. The respondent can have one of the following nine relationships with *each partner*:

- (1) currently married
- (2) currently cohabiting
- (3) currently pregnant with the romantic partner

- (4) currently dating
- (5) most recent
- (6) married to the partner before but the relationship is *not* current
- (7) cohabited with the partner before but the relationship is *not* current
- (8) pregnant with the romantic partner before but the relationship is *not* current
- (9) dated the partner before but the relationship is *not* current.

We mainly used two variables, H4TR13 and H4TR14, to construct a variable (STATUS1) that tells the relationship status *per partner*.

Based on the detailed information from variable STATUS1, we created the second flag variable (FLAG2) to tell the type of relationship data *each respondent* has. The type of data that the respondent has can fall into one of the five categories:

- (1) only has one current relationship
- (2) has multiple concurrent relationships
- (3) only had non-current relationship(s) (married, cohabited, or pregnant with the romantic partner)
- (4) only had relationship(s) identified by the respondent as “most recent”
- (5) had both non-current relationship(s) and relationship(s) defined by the respondent as “most recent”

We used this flag variable, FLAG2, from Section 16B to generate a new data file (Section16B\_AIDwise). In this data set, each respondent has one observation with one id, AID, with FLAG2 telling the type of relationship data each respondent has. We then used AID to merge two data files, Section 16A and Section16B\_AIDwise, together and created a new flag variable, FLAG3, which tells the type of relationship data each respondent has after the merge. Some respondents (total N=113) are in Section 16A but do not have data in Section 16B. Their relationship data have some missing in Section 16A questions. They are coded as 7 to 10 in FLAG3.

### Step 3.

This step meant to merge Section 16A, 16B, and 17 to construct the final current relationship status variable. We first briefly describe characteristics of Section 17.

Section 17 is about the respondent's relationship with one partner in detail. Only one relationship was automatically selected during the Wave IV Computerized Personal Interview. The survey used a set of criteria defined beforehand by the Add Health research team. The selected relationship was based on responses from Section 16A and Section 16B where the information about the relationship with each partner is available per respondent.

The criteria about the selection of the partner-specific relationship are described as follows: if the respondent has multiple current partners, the priority is given in this manner: marriage partner, cohabitation partner, pregnancy partner, dating partner. If two or more partners fall in the same type of relationship, the longer/longest relationship is selected. If two or more partners fall in the same type of relationship, and they are of the same duration, the respondent is asked to pick the partner they care about the most. If there are no current partners, the most recent partner is

selected. If there is no current partner or most recent partner, end dates for each marriage, cohabitation, and relationship with a pregnancy are reviewed to select the partner with the most recent end date. If two or more partners have the same end date, the one with longer/longest relationship is selected.

In summary, Section 17 only selected *one partner per respondent*. This section is different from Section 16B which includes information about multiple partners per respondent. The type of selected relationship in Section 17 can be a current one, non-current one, or “most recent” one.

#### Step 4.

Since our goal was to construct a current relationship status variable using information from Section 16A, 16B, and 17, we needed to take additional steps. We used AID and Partner ID (PTNR\_ID) to merge Section 17 with Section 16B (stacked relationship data file with possible multiple partners per respondent) which included from Step 2 two constructed variables, STATUS1 and FLAG2. Through this process, we were able to find out the partner-specific relationship that was selected into Section 17. Some respondents (total N= 23) have data in Section 16B but were not selected into Section 17. These cases were flagged in the fourth flag variable, FLAG4, and coded as 11 and 12. After the merge across Section 16A, 16B and 17, we used STATUS1 (partner-specific status variable) constructed from Section 16B and created a new relationship status variable (STATUS2). Respondents who were not selected into Section 17 (coded as 11 & 12 in FLAG4) and those who had missing data in Section 16A and 16B (coded as 7 to 10 in FLAG3 & FLAG4) were coded as missing in STATUS2.

Summary of Constructed Relationship Status Variables		
Section	Constructed Variable	Variables for Release
16A	FLAG1	N/A (intermediate variable)
16B	STATUS1	N/A (intermediate variable)
	FLAG2	N/A (intermediate variable)
16A & 16B_AIDwise*	FLAG3	C4VAR050
16A, 16B, 17 & 16B_AIDwise	FLAG4	C4VAR051
	STATUS2	C4VAR049
*Section 16B_AIDwise is a constructed data file.		

## Summary

Finally, we chose three variables for release, including one constructed relationship status variable, **C4VAR049** (i.e., STATUS2), and two flag variables, **C4VAR050** (i.e. FLAG3) and **C4VAR051** (i.e. FLAG4).

C4VAR049 (i.e., STATUS2) is able to tell the current relationship each respondent had by Wave IV. C4VAR050 (i.e. FLAG3) describes the type of relationship data each respondent has in the merged data file of Section 16A and 16B. C4VAR051 (i.e. FLAG4) illustrates the type of relationship data each respondent has when Section 16A, 16B, and 17 were merged together. FLAG1, STATUS1, and FLAG2 functioned as intermediate variables. They were discussed in the documentation to illustrate how data were processed and not necessary for released.