Add Health Relationship and Fertility Data

Bianka Reese, MSPH
Predoctoral Trainee
Carolina Population Center
University of North Carolina at Chapel Hill

June 21, 2016

Overview of Presentation

• Brief introduction: Relationship and fertility data in Add Health Waves I to III

• Wave IV:
  – Distinct listing of relationships in Wave IV
  – Nested data structure
  – Relationship and fertility data across sections
    • Content and criteria
    • Data flows
    • Hierarchical data relationships
    • Unique ID components
Overview of Presentation

• Computational and technical tips and example SAS code
  – Unique ID Creation
  – Summary counts
  – Different levels and units of analysis
  – Hierarchical file merging
• Common issues when working with relationship and fertility data
• Some general advice
• Wave V preview

Q & A
<table>
<thead>
<tr>
<th>Waves I, II</th>
<th>Wave III</th>
<th>Wave IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Demographic</td>
<td>• Demographic</td>
<td>• Demographic</td>
</tr>
<tr>
<td>• Family, siblings, friends</td>
<td>• Family, siblings, friends</td>
<td>• Family, siblings, friends</td>
</tr>
<tr>
<td>• Education, work</td>
<td>• Education, work, military</td>
<td>• Education, work, military</td>
</tr>
<tr>
<td>• Physical and mental health</td>
<td>• Physical and mental health</td>
<td>• Physical and mental health</td>
</tr>
<tr>
<td>• Daily activities and sleep</td>
<td>• Daily activities and sleep</td>
<td>• Daily activities and sleep</td>
</tr>
<tr>
<td>• Relationships</td>
<td>• Relationships</td>
<td>• Relationships</td>
</tr>
<tr>
<td>• Sexual, &amp; fertility histories</td>
<td>• Sexual, &amp; fertility histories</td>
<td>• Sexual, &amp; fertility histories</td>
</tr>
<tr>
<td>• Substance use</td>
<td>• Substance use</td>
<td>• Substance use</td>
</tr>
<tr>
<td>• Delinquency and violence</td>
<td>• Involvement w/criminal justice system</td>
<td>• Involvement w/criminal justice system</td>
</tr>
<tr>
<td>• Attitudes, religion</td>
<td>• Attitudes, religion</td>
<td>• Work attitudes and chars, religion</td>
</tr>
<tr>
<td>• Economics, expectations</td>
<td>• Economics, expectations</td>
<td>• Economics, expectations</td>
</tr>
<tr>
<td>• Psychological, personality</td>
<td>• Psychological, personality</td>
<td>• Big 5 Personality, stressors</td>
</tr>
<tr>
<td></td>
<td>• Children and parenting</td>
<td>• Children and parenting</td>
</tr>
<tr>
<td></td>
<td>• Civic participation</td>
<td>• Civic participation</td>
</tr>
<tr>
<td></td>
<td>• Gambling</td>
<td>• Cognitive function</td>
</tr>
<tr>
<td></td>
<td>• mentoring</td>
<td>• Psychosocial factors</td>
</tr>
</tbody>
</table>
Wave I & Wave II

Wave I (1994-1995; Grades 7-12; N=20,745)
Wave II (1996; Grades 8-12; N=14,738)

- Adolescents name up to 3 “romantic” relationships.
- If no romantic relationships, then name 3 “like” relationships

Information about relationships:

- Begin/end dates
- Social context when relationship began
- Partner demographics
  - Biological sex, age, race/ethnicity
- First & most recent sexual intercourse
  - date, contraceptive use
- Verbal and physical abuse (W2 only)
- Ideal relationship characteristics and experiences
Wave I & Wave II

**Wave I** (1994-1995; Grades 7-12; N=20,745)
**Wave II** (1996; Grades 8-12; N=14,738)

### Fertility Experiences
(female respondents only):

**Pregnancies:**
- Number of times, dates, outcome
- Contraceptive use, want pregnancy?
- Marriage at time of pregnancy
- Child alive?
- Child gave up for adoption?
Wave III (2001-2002; Aged 18-26; N=15,197)

Romantic and pregnancy relationships since 1995
(N = 42,330 relationship records)

Relationships in Detail
(N = 38,375)

3 sets selected according to different criteria:

- **Sexual relationships**, N = 36,128
- **2 most important relationships**, N = 20,878
- **Couple (CP) sample for partner recruitment**, N = 4,236
Couples Sample for partner recruitment
N=4,236

Sexual Relationships
N=36,128

2 Most important
N=20,878
Couples Sample

The couples sample was a purposive, quota sample (N=1,507 partners) designed to collect information on 1/3 married, 1/3 cohabiting, and 1/3 dating partners.

Relationships are eligible if:
- Current
- Opposite sex partners
- Partner 18 or older
- Relationship duration >= 3 months

Sampling weight: W3PTNR
### Information about relationships:

#### Marriage history
- Begin/end dates
- Current?
- Still living together – last date
- How ended?

#### Cohabitation history
- Begin/end dates
- Still living together – last date
- Ever married to partner?
- How ended?
Wave III

Relationships in Detail

R & partner’s age:
• Age when first met; began a romantic relationship; began sexual relationship

Begin/end dates:
• Romantic relationship; sexual relationship; marriage

Duration:
• Living together before marriage; sexual relationship before cohabitation; same residence

Cohabitation:
• Travel distance; other residence; who else lived there; # nights together
Wave III

Relationships in Detail (cont.)

Various types of sexual activities
- When: first/most recent time – dates, frequencies
- Enjoyment of sexual activity with partner

Different types of birth control/condom use
- Combined with different types of sexual activities
- When used: first/most recent time – dates, frequencies

Relationship qualities
- Best description of current relationship (dating exclusively?)
- Currently engaged to marry?
- Abuses (verbal, physical, forced sex, resulted injuries)
- Satisfaction; love; dominance
- Joint account; purchase; chores
WAVE IV RELATIONSHIP AND FERTILITY DATA
Wave IV (2007-2008; Aged 24-32; N=15,701)

Relationship data collection structure differs from Wave III:

- Specific romantic/sexual relationships R have had
- Exclusive listings by partner types:
  - Married (‘M’)
  - Cohabiting (‘C’)
  - Pregnancy (‘P’)
  - Currently Dating (‘CD’)
  - Most Recent (‘MR’)

Note: You cannot update Wave III with Wave IV relationships -- It's not an add-on!
Nested Data Structure:

- Respondent’s reporting on **total number of relationships/partners** (by types)
  - Section 16A [H4TR*]
    - Married (‘M’)
    - Cohabiting (‘C’)
    - Pregnancy (‘P’)
    - Currently Dating (‘CD’)
    - Most Recent (‘MR’)

- **Time-segments** associated with each relationship/partner
  - Section 16B & 16C [H4TR*]

- Selected relationship in **detail**
  - Section 17 [H4RD*]
Nested Data Structure (cont.):

- **All pregnancies** associated with the reported relationship/partners
  - Section 18 [H4PG*]

- **All live births** resulting from pregnancies
  - Section 19 [H4LB*]

- **All live births** who are still living and not given up for adoption (children)
  - Section 20A [H4KK*]

- Respondents who have at least 1 child at home reporting on their parenting attitudes.
  - Section 20B [H4KK*]
Wave IV: Romantic/Sexual Relationships

S16A: R’s description of total no. of relationships, pregnancies, live births, children still living.

- **Unit of Analysis** – Respondent \( (N=15,701) \)
- **Record ID:** AID
- **Contents:**
  - # Persons ever married (‘M’ partners)
  - # Persons ever lived with but not married (‘C’ partners)
  - # Persons ever had a pregnancy with besides ‘M’ and ‘C’ partners (‘P’ partners)
  - # Persons currently dating besides ‘M’, ‘C’, and ‘P’ partners (‘CD’ partners)
  - # Other relationships < 6 months since 2001
  - # Other relationships 6+ months since 2001
Wave IV: Romantic/Sexual Relationships

S16A: R’s description of total no. of relationships, pregnancies, live births, children still living.

- **Unit of Analysis** – **Respondent** (N=15,701)
- **Record ID:** AID
- **Contents:**
  - Currently pregnant? {for female respondents}
  - Total # (life-time) pregnancies (including current)
  - Total # (life-time) live births
  - Total # biological children still living? {for respondents and any partners}
  - Ever have trouble getting pregnant or avoiding a miscarriage? {for respondents and any partners}
Section 16A: Relationships – Individual Level

START S16A

H4TR1 (# M partners)

H4TR2 (# C partners)

H4TR3 (# P partners)

H4TR4 (# CD partners)

H4TR5 (# Romantic/sex partners lasted 6+ months since 2001)

H4TR6 (# Romantic/sex partners lasted < 6 months since 2001)

Female

H4TR7 (Currently preg?)

R’s Biological Sex

Male

Else

H4TR9 (# Times preg/impregn)

Else

H4TR8 (Probably preg?)

DK, RE

H4TR10 (# Live births)

Else

H4TR11 (# Children still living)

0, DK, RE

0, DK, RE

Else

H4TR12 (Trouble get pregn/impregn?)

End S16A
Wave IV  Hierarchical Data Structure of Individual, Relationship, Pregnancy, Live Birth, and Child Records

At Individual Level:
Obtain no. of different types of partners

No. of person(s) ever married

Besides

No. of person(s) ever lived with

Besides

No. of person(s) resulted in a pregnancy

Besides

No. of person(s) in current relationships

Besides

No. of person(s) romantically/sexually involved since 2001 (6+ or < 6 months)

At Relationship Level:
Generate relationship records: Obtain initials of each partner

M partners

Besides

C partners

Besides

P partners

Besides

CD partners

MR partner
**S16B: Relationship and partner characteristics**

- **Unit of Analysis** – Partners (N=30,263)
- **Record ID:** AID + PTNR_ID
- **Contents:**
  - Relationship current?
  - # Times married to partner (‘M’)
  - Ever lived together? (‘M’)
  - # Times lived together with partner (‘M’, ‘C’)
  - # Times pregnant with partner (‘M’, ‘C’, ‘P’)
  - Partner demographics: biological sex, age, Hispanic ethnicity, and race.
Section 16B: Relationships – Partner Level

2. Evaluate answers H4TR1-H4TR6.
   - If all are 0 (H4TR1 = 0, H4TR2 = 0, H4TR3 = 0, H4TR4 = 0, H4TR5 = 0, H4TR6 = 0), go to End S16C.
3. Otherwise, loop through all partners (S16B and S16C).
4. For each partner:
   - H4TR15: (# Times married to partner)
   - H4TR16: (Ever lived w/ partner?)
     - Yes: H4TR17: (# Times lived w/ partner)
     - Else: Collect Demographic Information (H4TR19 to H4TR24) for Each M, C, P, CD, MR Partner.
   - CD, MR Partners
   - H4TR13: (Fill in relationship type for each partner)
Collect Demographic Information (H4TR19 to H4TR24) for Each M, C, P, CD, MR Partner

H4TR19
(Partner’s sex)

H4TR20
(Partner’s current age)

H4TR21
(Partner older/same/younger as R?)

H4TR22
(# Years older/younger)

H4TR23
(Partner Hispanic/Latino?)

H4TR24
(Partner’s race)

Loop thru all partners: H4TR14 to H4TR24
S16B: Relationship and partner characteristics

Among 15,701 respondents, 2.5% did not report any partner records.
**S16C: Relationship time-segments**

- **Unit of Analysis** – *Time-segment* (N=36,133)
- **Record ID:** AID + PTNR_ID + H4TR25 + H4TR26
- **Contents:**
  - Time-segment number \([H4TR26]\)
  - Relationship begin and end dates
  - Best estimate of relationship duration
  - How marriage ended

Wave IV: Romantic/Sexual Relationships
Wave IV Hierarchical Data Structure: Relationship and Relationship Time-Segments

- Loop through all partners
  - M partners
    - How many times married to partner?
    - How many times live with partner?
  - C partners
    - How many times live with partner?
  - P partners
  - CD partners
  - MR partner

Relationship Time-Segments
- Begin/end Dates, Duration
- 1st 2nd Nth time...
- >0 times
### Wave IV: Romantic/Sexual Relationships

**S16C: Relationship time-segments (example from 1 respondent)**

<table>
<thead>
<tr>
<th>Partner ID</th>
<th>Partner Type</th>
<th>Relationship Time-Segment Type</th>
<th>Time-Segment Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTNR_ID</td>
<td>H4TR13</td>
<td>H4TR25</td>
<td>H4TR26</td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>M (1st time)</td>
<td>1=1st time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M (2nd time)</td>
<td>2=2nd time</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>M (1st time)</td>
<td>1=1st time</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>C</td>
<td>8=All</td>
</tr>
<tr>
<td>4</td>
<td>P</td>
<td>P</td>
<td>8=All</td>
</tr>
<tr>
<td>5</td>
<td>CD</td>
<td>CD</td>
<td>8=All</td>
</tr>
</tbody>
</table>
Section 16C: Relationships – Time-segment Level

- **H4TR25**
  - Fill in relationship segment type for each segment associated with each partner

- **H4TR26**
  - Fill in relationship segment type number for each segment

- **H4TR27M**
  - (Relationship began: M/Y)

- **H4TR27Y**
  - (Relationship began: M/Y)

- **H4TR28M**
  - (Relationship ended: M/Y)

- **H4TR28Y**
  - (Relationship ended: M/Y)

- **TSDURATN**
  - (Best estimate of relationship duration in months)

- **H4TR25 = marriage segment and H4TR28 M/Y NE current**

  - True
    - Loop to next segment of same partner, if applicable
  - Else
    - Loop thru all partners: H4TR27M to H4TR29

- **H4TR29**
  - (How marriage ended?)

- **End S16C**
S16B: Relationship and partner characteristics

Among 30,263 partner records, 81% link to 1 time-segment; 19% link to 2 time-segments; only a few linked to 3 or more time-segments.
Wave IV: Romantic/Sexual Relationships

S17: Relationship in detail

- **Unit of Analysis** – **Partner** (N=15,216)
- **Record ID:** AID + PTNR_ID
- **Contents:**
  - Age when relationship began
  - Duration of relationship
  - Relationship description and based on relationship:
    - Either partner have a home other than the one cohabiting?
    - # nights spent together
    - Living situation for spouse
  - Relationship qualities (enjoy ordinary things, handling problems, finances, partner listens, sex life, trust partner, love for partner, happy w/ partner, commitment, closeness)
S17: Relationship in detail

Contents (cont.):

- Chance marrying partner
- Likelihood relationship permanency
- Frequency of sexual activities
- Contraceptive use
- On both R’s and Partner’s side:
  - Concurrent partner?
  - Any violence?
  - Slap/hit/kick?
  - Injured?
  - Forced sex?
Criteria to Choose 1 Partner among All Reported For Gathering Detailed information in S17

START

Any CURRENT relationship?

No

Select 1

Care most

=1

R select partner whom CARE most

all same

Compare durations

longest

Select

Yes

Any MARRIED partner?

No

Any COHABITATN partner?

Yes

Any PREGNANCY partner?

No

CURRENT DATING partner

=1

How many?

> 1

How many?

No

Yes

Select

Select

Select

Select

Select

Select

How many?

> 1

How many?

all same

Compare END dates

most recent

Compare durations

longest

all same

MOSR RECENT partner?

No

How many other partners?

> 1

Yes

Select

Select

Select

Select
Section 17: Relationship in Detail

START S17

H4TR13
H4TR14
(from S16B)

Else

H4RD1
(R’s age when relationship began)

H4RD2D/M/Y
(Duration involved w/ partner)

H4TR13
(P, CD, MR partner)

C partner

H4RD3
(Relationship description)

H4RD4
(Residence other than cohab?)

Yes, DK, RE

H4RD5
(# Night/week together)

H4RD6
(Current status of marriage to partner)

M partner

Measure Relationship Qualities

H4TR13 = CD partner OR H4TR14 = Yes (Relationship current)
Among 15,315 respondents who have at least 1 partner record in S16B, 15,216 of them gave details on 1 partner in S17.
Wave IV: Fertility History

S18: Pregnancy Table

- **Unit of Analysis** – Pregnancy (N=21,966)
- **Record ID**: AID + PTNR_ID + PRGNO
- **Contents**:
  - Pregnancy outcomes
    - Abortion
    - Ectopic/tubal pregnancy
    - Miscarriage
    - Still birth
    - Live birth (C-section)
    - Live birth (vaginal delivery)
    - Pregnancy not ended
S18: Pregnancy Table

- Pregnancy end/due dates
- # live birth(s) per pregnancy
  - Babies fraternal/identical?
- Birth control use prior to pregnancy
- Want child then?
- Relationship context at pregnancy/birth
- R/partner went to prenatal care together;
- # weeks pregnant at 1st visit (female Rs only)
- Drinking and smoking habits during pregnancy (female Rs only)
Wave IV Hierarchical Data Structure: Relationship and Pregnancy Records

- M partners
  - How many times got/made pregnant by partner?
    - 1st reported
    - 2nd reported
    - Nth reported...
    - ...

- C partners
  - How many times got/made pregnant by partner?
    - 1st reported
    - 2nd reported
    - Nth reported...
    - ...

- P partners
  - How many times got/made pregnant by partner?
    - 1st reported
    - 2nd reported
    - Nth reported...
    - ...

- CD partners
- MR partner

Loop through M, C, P partners only
Wave IV: Fertility History

S18: Pregnancy Table

Among 15,701 respondents, 40% reported no pregnancy record in S18; 20% reported one pregnancy record; 19% reported two; 11% reported three; 10% reported 4 or more
Among 21,196 pregnancy records, 65% live birth; 28% abortion/miscarriage; 4% current pregnancy; 3% other
Section 18: Pregnancy Table (Cont.)

Collect Relationship Context during Pregnancy: H4PG9 to H4PG11

- **Marriage partner**
  - H4TR13 Relationship Type
    - Cohabitation partner
      - H4PG9 (Married ?)
        - No, DK, RE
          - H4PG10 (Living together ?)
          - No, DK, RE
            - H4PG11 (Relationship context)
          - Yes
            - Collect Prenatal Care and Habits during Pregnancy: H4PG12 to H4PG15
        - Yes
          - Collect Prenatal Care and Habits during Pregnancy: H4PG12 to H4PG15
Section 18: Pregnancy Table (Cont.)

Collect Prenatal Care and Habits during Pregnancy: H4PG12 to H4PG15

C-Section, vaginal delivery, currently pregnant

H4PG1 Pregnancy outcome

Else

H4PG12 (Prenatal care?)

Female AND H4PG12 = yes

R's Biological Sex

Male

Else

H4PG13 (Week preg 1st prenatal visit)

C-section or vaginal delivery

H4PG14 (Smoke: freq)

H4PG15 (Drink: freq)

Loop to next pregnancy of same partner, if exists

Loop thru all relevant partners: H4PG1 to H4PG15

End S18
S19: Live Births

- **Unit of Analysis**: Live birth (N=14,749)
- **Record ID**: AID + PTNR_ID + PRGNO + LBNO
- **Contents**:
  - Confirm baby’s birth date; corrected birth date
  - Baby’s biological sex
  - Baby went home; Why not?
  - Birth weight
  - Born before/on/after due date? # weeks/days before/after
  - Still living; When died?
Wave IV Hierarchical Data Structure: Pregnancy and Live Birth Records

Pregnancy Records:
Collected for Rs who have had at least 1 pregnancy (including current)
ID components: R’s ID + Partner’s ID + Pregnancy no.

For Each Pregnancy, ask:
Pregnancy Outcome, Pregnancy End/Due Date
No. Live Birth per Pregnancy, Relationship context with partner
Prenatal Care, Health Risk Habit (smoke, alcohol)

Generate Live Birth Records
Check and correct baby’s birthday (i.e., Pregnancy End Date)
Gather baby’s initial/name

If sum of live births (for all pregnancies) per R = 0,
then skip Live Birth Section

If NO step/adopted/foster child reported in Household Roster, then skip Child and Parenting Section

Live Birth Records:
Collected for Rs who have had at least 1 live birth
ID components: R’s ID + Partner’s ID + Pregnancy no + LB no
Section 19: Live Births

START S19

Total no. of pregnancy with live birth outcome (from S18)

Else

Loop thru all pregnancy with live birth outcome

Pregnancy end date (from S18)

End S19

= 0

H4LB1 (BB birth date correct?)

Yes

H4LB3 (BB’s sex)

H4LB4 (BB went home?)

Yes, DK, RE

No

H4LB5 (Why not?)

Collect Baby’s Birth Weight, Height and Other Characteristics: H4LB6 to H4LB11
Section 19: Live Births (Cont.)

Collect Baby’s Birth Weight, Height and Other Characteristics

\[H4LB6P/O\] (Birth weight)
- < 5 lb 8 oz, DK, RE

\[H4LB7\] (BB weight < 5.5 lb)
- Before, After

\[H4LB8\] (Born bef/aft/on due date)
- Died in hospital

\[H4LB9W/D\] (Week/day bef/aft due date)
- Else

\[H4LB5\]
- Else

\[H4LB10\] (BB still living?)
- No

\[H4LB11M/Y\] (BB died month/year)
- Loop thru preg w/ LB outcome among all partners
- End S19

Yes, DK, RE
- Loop to next LB of same pregnancy, if exists
Among 15,701 respondents, 51% reported no live birth record in S19; 21% reported one record; 17% reported two; 7% reported three; 4% reported 4 or more.
S20: Children and Parenting, Part A

- **Unit of Analysis** – Child (N=14,553)
- **Record ID:** AID + PTNR_ID + PRGNO + LBNO
- **Contents:**
  - How old?
  - Child lives with respondent; Ever?
    - Last date living together
    - With whom child lives with
    - Distance
    - How often R see child?
  - Respondent lives with partner?
    - If not, distance?
    - Any custody agreement?
    - How often partner see child?
  - Child’s general health
  - Child-specific health problems
S20: Children and Parenting, Part B

- **Unit of Analysis** – Respondent (N=14,553)
- **Record ID:** AID
- **Contents:**
  - Language speak to child(ren) at home?
  - Happy in role as parent?
  - Respondent feels close to his/her child(ren)?
  - Child(ren): major source of stress?
  - Respondent feels overwhelmed by responsibility of being a parent
Wave IV Hierarchical Data Structure:
Live Birth and Child Records

Live Birth Records, ask:
BB birth weight, height, Gender, BB still alive?

If sum of (alive BB – given up) per R = 0 → true

Check Household Roster for any Step/adopted/foster Child

Generate Child Records

If > 0

Child Records:
Collected for Rs who have had at least 1 alive baby not given up for adoption
ID Components: R’s ID + Partner’s ID + Pregnancy no. + LB no.

For each child, ask:
Age, Living arrangement (family structure), Child custody, Health conditions

If > 0

Loop through all child records per R

At Individual Level, ask:
Language use at home, Parenting experiences
Section 20 Part B: Individual level -- Parenting

1. Collect R’s Parenting Attitudes

2. Check if the total number of LBs still living (from S19) minus the total number of children given up for adoption (sum of H4KK6H=1) is greater than 0.
   - If True, proceed to H4KK14 (Language spoken to child at home).
   - If False, proceed to step 3.

3. If the total number of step, foster, and adopted children (from S8) is greater than 0, proceed to the end.
   - If False, proceed to H4KK15A (R happy as parent).

4. H4KK15A (R happy as parent)
   - If True, proceed to H4KK15B (R feels close to child(ren)).
   - If False, proceed to H4KK15C (Child(ren) major stress to R).

5. H4KK15C (Child(ren) major stress to R)
   - If True, proceed to H4KK15D (Parent responsibility overwhelming).
   - If False, proceed to the end.

6. H4KK15D (Parent responsibility overwhelming)
   - If True, proceed to the end.
   - If False, proceed to the end.

7. End S20
Among 15,701 respondents, 51% reported no child record in S20A; 21% reported one child record; 17% reported two; 7% reported three; 3% reported 4 or more.
Wave IV Hierarchical Data Structure:
Pregnancy, Live Birth, and Child Records

Partner Level

Pregnancy Level

Live Birth Level

Child Level

Partner #1
(M Partner)

Partner #2
(M Partner)

Partner #3
(C Partner)

Partner #4
(P Partner)

Pregnancy #1

Pregnancy #2

Pregnancy #1

Pregnancy #1

Abortion

Plural Live Births

Current Pregnancy

Live Birth #1

Live Birth #2

Live Birth #1

Live Birth #1

Child #2

Child #2

Child #1

Child #1

Still Birth
COMPUTATIONAL AND TECHNICAL TIPS
Data Structure Transformation for Hierarchical Files

1. Construct unique record IDs for each file

   - **Individual ID:** $AID + PTNR\_ID + PRGNO + LBNO$
   - **Relationship ID:** $AID + PTNR\_ID$
   - **Time-segment ID:** $AID + PTNR\_ID + H4TR25 + H4TR26$
   - **Pregnancy ID:** $AID + PTNR\_ID + PRGNO + LBNO$
   - **Live birth ID:** $AID + PTNR\_ID + PRGNO + LBNO$
   - **Child ID:** $AID + PTNR\_ID + PRGNO + LBNO$
Data Structure Transformation for Hierarchical Files

2. Lower-level file contains IDs of all higher-level files

- Individual ID
- Relationship ID
- Pregnancy ID
- Live birth ID

For example:

Pregnancy records contain record IDs of the corresponding individual, relationship, and pregnancy record ID.
Example SAS code

*---------------------------------------------------------------*
| Construct partner record ID and |
| pregnancy record ID (Character vars) | | *
*---------------------------------------------------------------*

data pregrcds (drop=nreln npreg);
  length partner_ID $10  preg_ID $12  nreln npreg $2;

  set pregtable (keep=AID PTNR_ID PRGNO ...);
  if PTNR_ID <=9 then nreln="0"||put(PTNR_ID,1.);
  else nreln= PTNR_ID;
  if PRGNO <=9 then npreg="0"||put(PRGNO,1.);
  else npreg= PRGNO;
  preg_ID =AID||nreln||npreg;
  partner_ID =AID||nreln;

  label preg_ID="Preg Record ID (AID+PTNR_ID+PRGNO)"
    partner_ID="Partner Record ID (AID+PTNR_ID)"

; run;
Compute Summary Counts

1. Describe respondent by his/her reporting on total numbers of ‘event’: e.g.,
   - Total no. of partners reported by R
   - Total no. of pregnancies reported by R

2. Describe respondent by his/her average reporting on a certain ‘event’: e.g.,
   - Average no. of pregnancy per partner by R
Example SAS code

*---------------------------------------------------------------*
| Total no. of partner records reported by R.                   |
*---------------------------------------------------------------*;
proc sort data=partnrcds; by AID; run;
proc summary data=partnrcds;
  var partnercd;
  by AID;
  output out=sumreltn sum(partnercd)=tot_partnr;
run;

*---------------------------------------------------------------*
| Total no. of pregnancy records reported by R.                |
*---------------------------------------------------------------*;
proc sort data=pregnrcds; by AID; run;
proc summary data=pregnrcds;
  var pregncd;
  by AID;
  output out=sumpreg sum(pregrcd)=tot_preg;
run;

For syntax, visit Add Health website here
Example SAS code

*---------------------------------------------------------------------*
| Average no. of pregnancy per partner reported by R. |
*---------------------------------------------------------------------*;
proc sort data=individ; by AID; run;
data individ2;
  merge individ (in=one)
    sumreltn (in=two keep=AID tot_partnr)
    sumpreg (in=three keep=AID tot_preg);
by AID;
  if one then W4main=1; else W4main=0;
  if two then rpt_partnr=1; else rpt_partnr=0;
  if three then rpt_preg=1; else rpt_preg=0;
  label tot_partnr="Total no. of partner rcrds reported by R"
  tot_preg  ="Total no. of preg rcrds reported by R"
  rpt_partnr="Rs who reprted at least 1 partner record"
  rpt_preg  ="Rs who reprted at least 1 preg record"
  W4main    ="Rs who are in W4 Main file"
;
Example SAS code

*--------------------------------------------------------------- *
| Assign '0' to missing values in total no. of partner records and in total no. of pregnancy records reported by R. |
*--------------------------------------------------------------- *
array ms1[*] tot_partnr tot_preg;
  do i=1 to dim(ms1);
    if ms1[i]=. then ms1[i]=0;
  end;
*--------------------------------------------------------------- *
| Compute average no. of pregnancy per partner reported by R. |
*--------------------------------------------------------------- *
  if tot_partnr ne 0 then do;
    ave_preg_partnr = tot_preg/tot_partnr;
  end;
  label ave_preg_partnr="R's average no. of pregnancy per partner";
run;

For syntax, visit Add Health website here
Example SAS code: Construction of Relationship Quality Measure

```sas
/* Current PRIMARY relationship quality */
if h4rd7a in (96,98) then h4rd7a2=.; else h4rd7a2=h4rd7a;
if h4rd7b in (96,98) then h4rd7b2=.; else h4rd7b2=h4rd7b;
if h4rd7c in (96,98) then h4rd7c2=.; else h4rd7c2=h4rd7c;
if h4rd7d in (96,98) then h4rd7d2=.; else h4rd7d2=h4rd7d;
if h4rd7e in (96,98) then h4rd7e2=.; else h4rd7e2=h4rd7e;
if h4rd7f in (96,98) then h4rd7f2=.; else h4rd7f2=h4rd7f;
if h4rd7g in (96,98) then h4rd7g2=.; else h4rd7g2=h4rd7g;
if h4rd7a2>. and h4rd7b2>. and h4rd7c2>. and h4rd7d2>. and h4rd7e2>. and h4rd7f2>. and h4rd7g2>. then do;
relnqual=sum(h4rd7a2,h4rd7b2,h4rd7c2,h4rd7d2,h4rd7e2,h4rd7f2,h4rd7g2);
end;
if currreln4cat in (1,2,3) then currrelnqual=relnqual;
else if currreln4cat=. then currrelnqual=.;
label currrelnqual='CurrRelnQualityScale7-35LowIsHappy';

/* Current PRIMARY relationship duration */
if currreln4cat in (1,2,3) then do;
currrelndurn=h4rd2y+(h4rd2m/12)+(h4rd2d/365.25);
end;
else if currreln4cat in (.,0) then currrelndurn=.;
label currrelndurn='DurationOfCurrRelnInYears(incDecimals)';
```

Add Health
The National Longitudinal Study of Adolescent to Adult Health
### Computational and Technical Tips

#### Examples for Different Levels and Units of Analysis, and Hierarchical Files Merging

<table>
<thead>
<tr>
<th>Level of analysis</th>
<th>Purpose</th>
<th>Input files</th>
<th>Data sorted by</th>
<th>Summary statistics/ descriptive per unit of analysis</th>
<th>Files merged by</th>
<th>Output file</th>
<th>Unit of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respondent</strong></td>
<td>Describe respondent by a summary of partner characteristics. e.g., Total no. of partners reported by respondent</td>
<td>Respondent AID</td>
<td>AID</td>
<td>Per respondent e.g., Count of partners</td>
<td>AID</td>
<td>Respondent</td>
<td><strong>Respondent</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Partner AID, PTNR_ID</td>
<td>AID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Respondent</strong></td>
<td>Describe respondent by a summary of pregnancy history. e.g., Total no. of pregnancies with abortion as outcome</td>
<td>Respondent AID</td>
<td>AID</td>
<td>Per respondent e.g., Count of pregnancies ended in abortion</td>
<td>AID</td>
<td>Respondent</td>
<td><strong>Respondent</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pregnancy AID, PTNR_ID, PRGNO</td>
<td>AID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of analysis</td>
<td>Purpose</td>
<td>Input files</td>
<td>Data sorted by for computation or merging</td>
<td>Summary statistics/descriptive per unit of analysis</td>
<td>Files merged by</td>
<td>Output file</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Respondent</td>
<td>Describe respondent by a summary of live birth history. e.g., Total no. of live births born outside marriage</td>
<td>Respondent, AID</td>
<td>AID</td>
<td>Per respondent e.g., Flag live births born while R and partner not married and count up total</td>
<td>AID</td>
<td>Respondent</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Live birth, AID, PTNR_ID, PRGNO, LBNO</td>
<td>AID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partner</td>
<td>Associate respondent's characteristics to partners. e.g., Identify hetero/homosexual relationship</td>
<td>Partner, AID, PTNR_ID</td>
<td>AID</td>
<td>Spread R’s information onto each partner record e.g., spread R’s biological sex onto partner record and compare</td>
<td>AID</td>
<td>Partner</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Respondent, AID</td>
<td>AID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example for Different Levels and Units of Analysis, and Hierarchical Files Merging (Cont.)**
### Computational and Technical Tips

#### Examples for Different Levels and Units of Analysis, and Hierarchical Files Merging (Cont.)

<table>
<thead>
<tr>
<th>Level of analysis</th>
<th>Purpose</th>
<th>Input files</th>
<th>Data sorted by for computation or merging</th>
<th>Summary statistics/descriptive per unit of analysis</th>
<th>Files merged by</th>
<th>Output file</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner</td>
<td>Associate pregnancy characteristics to partners. e.g., No. of non-birth pregnancies associated with each partner</td>
<td>Partner AID, PTNR_ID</td>
<td>AID, PTNR_ID</td>
<td>Per partner e.g., Flag pregnancies with non-birth outcomes and count up total.</td>
<td>AID, PTNR_ID</td>
<td>Partner</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pregnancy</td>
<td>AID, PTNR_ID, PRGNO</td>
<td>AID, PTNR_ID</td>
<td>By partner</td>
<td>AID, PTNR_ID</td>
</tr>
<tr>
<td>Partner</td>
<td>Associate partner demographics w/ partner detailed information in S17.</td>
<td>Partner AID, PTNR_ID</td>
<td>AID, PTNR_ID</td>
<td></td>
<td>AID, PTNR_ID</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Partner</td>
<td>AID, PTNR_ID</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Issues When Working with Relationship and Fertility Data

• Retrospective reports on relationship and fertility history
• Ages of first sexual experiences reported in whole years
• Sexual behavior and fertility during adolescence are sensitive topics that may be underreported
• Data items vary by wave
• Challenging programming jobs
• **Spend time to know your data!**
  – Examine question branching; Test skip patterns; Check value ranges; Further clean your data for odd/extreme cases. Know your variables!
  – Check your programs: When in doubt, list your cases and examine.
  – Do logical checks and analytical checks on your results. You can never be too careful!

• **Take small steps: Think through your results**
  – Do thorough preliminary analysis before going into sophisticated multivariate advanced statistics.
  – Take pains to document your data and programming decisions.
General Advice

• Do not assume!
  – Read as much as you can including:
    • Add Health documentation
    • Subscribe to Add Health list server:
      – Email Add Health (addhealth@unc.edu) and request that your email address be added.
    • Check out publications using Add Health data
<table>
<thead>
<tr>
<th>Relationships</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Marriage and cohabitation summary history</td>
</tr>
<tr>
<td></td>
<td>• Current/most recent partner demographics</td>
</tr>
<tr>
<td></td>
<td>• Gender, age, race/ethnicity</td>
</tr>
<tr>
<td></td>
<td>• Relationship duration and quality</td>
</tr>
<tr>
<td></td>
<td>• Frequency of sexual activity</td>
</tr>
<tr>
<td></td>
<td>• Frequency of contraceptive use</td>
</tr>
<tr>
<td></td>
<td>• Fertility intentions</td>
</tr>
<tr>
<td></td>
<td>• Concurrent partners</td>
</tr>
<tr>
<td></td>
<td>• IPV and injuries</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pregnancy and birth history</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Full history</td>
</tr>
<tr>
<td></td>
<td>• Pregnancy complications</td>
</tr>
<tr>
<td></td>
<td>• Pregnancy outcomes</td>
</tr>
<tr>
<td></td>
<td>• Birth outcomes</td>
</tr>
<tr>
<td></td>
<td>• Relationship with partner at time of child’s birth</td>
</tr>
<tr>
<td></td>
<td>• Child’s health</td>
</tr>
</tbody>
</table>
Thank you!

Questions?

Contact:
Bianka Reese
bmreese@live.unc.edu

For syntax and other resources, visit:
https://www.cpc.unc.edu/projects/addhealth/