The National Longitudinal Study of Adolescent Health

Using Secondary Data for Analysis of Marriage and Family

2009 ICPSR Summer Program Workshop
Ann Arbor, Michigan

Mariah Mantsun Cheng, Phd.
Carolina Population Center, UNC-Chapel Hill
Overview of Presentation

• Introduction to Add Health
  – Research Design and Study Components.
  – Add Health Website, Documentation, and Data Access.
Overview of Presentation

• Questionnaires Focus on Marriage and Family
  – Crosswalk over Waves:
Overview of Presentation

- Questionnaires Focus on Marriage and Family (Cont.)
  - Nested Data Structure:
    - Individual
    - Relationship
    - Pregnancy
    - Live Birth
    - Child
Overview of Presentation (Cont.)

• Computational and Technical Tips
  – Data Structure Transformation:
    • From Hierarchical to Individual
  – Summary Count Computation
  – Some Data Cleaning and Programming Tips
Overview of Presentation (Cont.)

• Wave IV
  – Data Release
  – Some General Advice

• Questions and Answers
Add Health

- On-going program project that began in 1994.
- Developed in response to a congressional mandate to fund a study of adolescent health.
- Funded by the National Institute of Child Health and Human Development (NICHD) with co-funding from 17 other federal agencies.
- Fourth wave of data collection funded in 2006.
Key Features of Add Health

• Nationally representative study that explores the causes of health and health-related behaviors of adolescents and their outcomes in young adulthood.

• Multi-survey, multi-wave inter-disciplinary design.

• Direct measurement of the social contexts of adolescent life and their effects on health and health behavior.

• Unprecedented racial and ethnic diversity and genetically informed sibling samples.
Sampling Structure for Add Health

School Sampling Frame = QED

- HS
- Feeder

Sampling Frame of Adolescents and Parents  N = 100,000+ (100 to 4,000 per pair of schools)

Genetic Samples
- Saturation Samples from 16 Schools
- Disabled Sample

Main Sample 200/Community

Ethnic Samples
- High Educ Black
- Puerto Rican
- Chinese
- Cuban

Identical Twins
Fraternal Twins
Full Sibs
Half Sibs
Unrelated Pairs in Same HH
Genetic Supplements: Siblings

- **Twins**
  - Any student identified as a twin (in in-school survey)
  - Previously unreported twins discovered in in-home interviews.

- **Other siblings of twins**
  - Non-twin siblings of those in the twin sample (in Grades 7-12 at time of sample selection)

- **Other full siblings**
  - Neither member was a twin and both were in Grades 7-12

- **Half-siblings**
  - Pairs of half-siblings of which both were in Grades 7-12

- **Non-related**
  - 7-12 graders living in same household but did not share the same biological mom or dad.
<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>1,767</td>
<td>8.5</td>
</tr>
<tr>
<td>Cuba</td>
<td>508</td>
<td>2.5</td>
</tr>
<tr>
<td>Central-South America</td>
<td>647</td>
<td>3.1</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>570</td>
<td>2.8</td>
</tr>
<tr>
<td>China</td>
<td>341</td>
<td>1.7</td>
</tr>
<tr>
<td>Philippines</td>
<td>643</td>
<td>3.1</td>
</tr>
<tr>
<td>Other Asia</td>
<td>601</td>
<td>2.9</td>
</tr>
<tr>
<td>Black (Africa/Afro-Caribbean)</td>
<td>4,601</td>
<td>22.2</td>
</tr>
<tr>
<td>Non-Hispanic White (Eur/Canada)</td>
<td>10,760</td>
<td>52.0</td>
</tr>
<tr>
<td>Native American (non-Hispanic)</td>
<td>248</td>
<td>1.2</td>
</tr>
<tr>
<td><strong>Total N</strong></td>
<td><strong>20,686</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

Missing on race/ethnicity=59
Wave I
1994-1995

In-School Administration

School Admin 144

Wave II
1996

In-School Administration

School Admin 128

Wave III
2001-2002

In-HOME Administration

Adolescents in grades 7-12
20,745

Parent 17,670

Wave IV
2007-08

Adolescents in grades 8-12
14,738

Young Adults Aged 18-26
15,197

Adults Aged 24-32

In-School Administration

Partners 1,507

In-Home Administration

Students 90,118

School Admin 144

School Admin 128

Students 128

Students 128

Students 128

Students 128
Add Health Questionnaires

- **Wave I**  
  1994-95  
  School Administrator, N=144  
  In-School Adolescent, N=90,118  
  In-Home Adolescent, N=20,745 (80%)  
  In-Home Parent, N=17,670

- **Wave II**  
  1996  
  School Administrator, N=128  
  In-Home Adolescent, N=14,738 (90%)

- **Wave III**  
  2001-02  
  Partner, N=1507  
  In-Home Young Adult, N=15,197 (77%)

- **Wave IV**  
  2007-08  
  In-Home Young Adult, N=15,500 (80%)

*estimate only*
Social, Behavioral, and Biological Linkages Across the Life Course

Longitudinal View of Add Health

Adolescence → Young Adulthood

Wave I
(20,745)
1994-95
12-19

Wave II
(14,738)
1996
13-20

Wave III
(15,197)
2001-02
18-26

Wave IV
(15,500)
2008
24-32

* estimate only

Social Environmental Data:

school
family
romantic rel
neighborhd
community
peer

school
family
romantic rel
neighborhd
community
peer

college
family
romantic rel
neighborhd
community
peer

college
family
romantic rel
neighborhd
community
peer
Key Accomplishments of Add Health

- Data made available to more than 3,500 investigators for analyses
- 350+ grants awarded to analyze data
- 1,000+ peer-reviewed publications
- Research related to obesity, physical exercise, sleep, diet, substance use, mental health, reproductive health, sexual behavior, anti-social behavior, neighborhood effects, social relationships, parenting, family formation, health disparities, health care...
A Few Publication Examples

- **Marriage, Family**
Cohabitation, Romantic, Non-romantic Relationship

A Few Publication Examples (Cont.)

- **Fertility, Pregnancy**
Website, Documentation, Data Access

- **Main page:** [www.cpc.unc.edu/addhealth](http://www.cpc.unc.edu/addhealth)
  - **Study Design:** [www.cpc.unc.edu/projects/addhealth/design](http://www.cpc.unc.edu/projects/addhealth/design)
  - **Data:** [www.cpc.unc.edu/projects/addhealth/data](http://www.cpc.unc.edu/projects/addhealth/data)

- **Restricted-Use Data**
  - Data description
    - [www.cpc.unc.edu/projects/addhealth/data/restricteduse/datasets](http://www.cpc.unc.edu/projects/addhealth/data/restricteduse/datasets)

- **Using the Data**
  - **User Guides:** [www.cpc.unc.edu/projects/addhealth/data/using/guides](http://www.cpc.unc.edu/projects/addhealth/data/using/guides)
    - *Strategies to Perform a Design-Based Analysis Using the Add Health Data* (Chantala & Tabor, 1999)
    - *Guidelines for Analyzing Add Health Data* (Chantala, 2006)
    - *Constructing Weights to Use in Analyzing Pairs of Individuals from Add Health Data* (Chantala, 2001)
Website, Documentation, Data Access (Cont.)

• Main page: www.cpc.unc.edu/addhealth (Cont.)
  – Codebooks: www.cpc.unc.edu/projects/addhealth/codebooks
    • Section Lists: www.cpc.unc.edu/projects/addhealth/codebooks/sectionlists (with section names and the corresponding variable prefixes for each section)
    • Indexes: www.cpc.unc.edu/projects/addhealth/codebooks/indexes (a quick reference to questions asked during interviews and corresponding variables)
    • Waves I, II, III codebooks ...(with descriptions of interview questions, skipped patterns, value codes, variables names, and frequencies etc.)
  – Publications: www.cpc.unc.edu/projects/addhealth/pubs

• Data Access
  – Restricted-use contractual dataset
  – Public-use file
  – ICPSR (download and limited on-line analysis www.icpsr.umich.edu/cocoon/DSDR/DAS/21600.xml)
<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>Educ, work, military (records)</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 and abuse</td>
</tr>
<tr>
<td>Delinquency and violence</td>
<td>Involvmt w/criminal justice system</td>
<td>Involvmt w/criminal justice sys</td>
</tr>
<tr>
<td>Attitudes, religion</td>
<td>Attitudes, religion</td>
<td>Work attitudes and chars, relig</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>Children and parenting</td>
<td>Children and parenting</td>
<td>Children and parenting</td>
</tr>
<tr>
<td>Civic participation</td>
<td>Civic participation</td>
<td>Civic participation</td>
</tr>
<tr>
<td>Gambling</td>
<td>Gambling</td>
<td>Gambling</td>
</tr>
<tr>
<td>mentoring</td>
<td>mentoring</td>
<td>mentoring</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Coverage of Marriage and Family Data

<table>
<thead>
<tr>
<th>Household Roster</th>
<th>Parents Information</th>
<th>Relations with parents</th>
<th>Siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendship Network</td>
<td>Ideal Romantic Relationship/ Marital Attitudes</td>
<td>Romantic &amp; Non-romantic Relationships</td>
<td>Sexual experiences &amp; Behaviors</td>
</tr>
<tr>
<td>Birth controls &amp; contraceptives</td>
<td>Marriage</td>
<td>Cohabitation</td>
<td>Partners Information</td>
</tr>
<tr>
<td>Pregnancies</td>
<td>Live Births</td>
<td>Children</td>
<td>Parenting</td>
</tr>
</tbody>
</table>
Collect information on people living with R:
- How many?
- Male or Female?
- Relationship Type
- How old?

Other information on surrogate parents
- Who in R’s household acts in the place of a mom/dad to R?

W1 In-home interview also asked:
- {Member} always lived in same household? If not, for how many months/years have R and {member} lived in same household?
Household Roster across Waves (Cont.)

- Relationship Types:
  
  wife/husband, partner, son, daughter, brother, brother’s wife, sister, sister’s husband, sister’s partner, father, father’s wife, father’s partner, mother, mother’s husband, mother’s partner, father-in-law, mother-in-law, grandfather, grandmother, great-grandfather, great-grandmother, uncle, aunt, cousin, nephew, niece, other relative, other non-relative

- IF son/daughter: biological, step, adopted, foster, other

- IF brother/sister: full, half, step, adoptive, foster, other
Family Structure Constructed from W1 Household Roster

- Two biological parents
- Two adoptive parents
- Bio Mom, Step Dad
- Bio Dad, Step Mom
- Two step/foster parents
- Single Mom
- Single Dad
- Surrogate parent(s)
Cumulative Probabilities of First Nonmarital Birth by Family of Origin Family Structure. (N=5,884 women)

Parents Information

In-home Data on Resident Parents

- Resident Mom and Dad: W1/W2
  - How far in school did s/he go?
  - Born in the US; What country?
  - Kind of work; Work for pay last 12 months; How many hours; Work at or outside home; Receive welfare; Disabled?
  - At home when R leaves for or returned from school?
  - At home when R goes to bed?
  - Ever smoked cigarettes?
In-home data on Non-resident parents

- Non-resident Biological Mom and Dad: W1/W2; W3
  - Still living; How old was R when s/he died?
  - How far in school did s/he go?
  - Disabled?
  - Born in the US?
  - Ever smoked cigarettes?
Additional data from W1 Parent Survey

- A parent (mostly the Resident mom) was interviewed:
  - Biological sex; How old? Born in US?
  - Detailed race/ethnicity background; Age at 1st marriage.
  - Marital status: Currently dating? Current relationship (begin/end year).
  - Work situation; Religious background; Voluntary organization membership; Total income; Welfare.
Additional data from W1 Parent Survey (Cont.)

- A parent (mostly the Resident mom) was interviewed:
  (Cont.)
  - Physical health; Medical care; Drinking/smoking habits.
  - Descriptions and sense of neighborhood.
  - Participation in {R}’s school
    - Talked w/ teacher
    - Volunteered at school
  - Evaluation of the school
    - Good; safe; high priority in learning
Additional data from W1 Parent Survey (Cont.)

• Parent gave info on her/his current spouse/partner:
  – *Detailed race/ethnicity background*
  – *Biological sex; Religion; Schooling*
  – *Work situation; Disabled? Welfare?*
  – *Relationship (how happy, separation, argument/fight)*
  – *General health; Drinking habit.*
Additional data from W1 Parent Survey (Cont.)

• Parent gave info on the child {R}:
  – {R} ever lived w/ biological mom/dad?
  – How much support s/he paid for {R}?

• Parent’s relationship to {R}
  – Ever been away from {R} for 6 months+; When (age of {R})?
Relations with Parents

In-home Data on Resident Parents:

- Do parents let R to make decisions about: W1/W2
  - *Time be home weekend nights*
  - *Go to bed on week nights*
  - *People hang around with*
  - *What to wear, eat*
  - *What to watch on TV*
    - *How much*
    - *Which programs*
In-home Data on Resident Parents (Cont.):

- How R feels about Resident Mom and Dad: W1/W2; W3
  - Mom/Dad is warm, loving, encouraging R to be independent.
  - R satisfied w/ his/her relationship w/ Mom/Dad, communications.
  - Mom/Dad knows what is going on in R’s life.
  - R feels close to Mom/Dad.
In-home Data on Resident Parents (Cont.):

- Did R do things with Resident Mom and Dad (last 4 weeks):
  - Shopping
  - Played sport
  - Attended church
  - Talked about dating, school work, personal problems
  - Had an argument
  - Gone to movies, museums, concert, sport events
  - Worked on school project
  - Or none?
In-home data on Non-resident Biological Parents:

- Contact with Non-resident Mom and Dad: W1/W2; W3
  - R knows him/her; Still living?
  - Ever lived with her/him?
  - How old was R when s/he last lived with her/him?
  - In last 12 months: How often R stayed overnight, talked with, did things with her/him?
In-home data on Non-resident Biological Parents (Cont.):

- How R feels about Non-resident Mom and Dad: W1/W2; W3
  - How close?

- Did R do things with Non-resident Mom and Dad (last 4 weeks): W1/W2
  - Shopping, played sport, attended church, talked about dating, school work, personal problems, had an argument, gone to movies, museums, concert, sport events, worked on school project, or none?
Relations with Parents (Cont.)

Additional info from W1 Parent Survey:

• Parent’s (mostly the Resident Mom) relations with \{R\}:
  – Academic expectations on \{R\};
  – Talked with \{R\} about school work and activities?
  – Know \{R\}’s best friend (school, ever met in person, friend’s parents)?
  – How many parents of \{R\}’s friends talked to in last 4 weeks?
Additional info from W1 Parent Survey (Cont.):

- Parent’s (mostly the Resident Mom) relations with {R} (Cont.):
  - Get along well with {R}; Make decision together; Understand, trust {R}; Feel close?
  - Discuss w/ {R} on birth control, sex, STD?
  - Approve/disapprove {R} having sex in various situations; Ever recommend a birth control method to {R}?
  - Know about {R}’s social life, dating, kissed, having sex?
In-home W1, R selected and ranked cards describing activities that would occur in an ideal romantic relationship:

- Go out together in a group
- Meet partner’s parents
- Tell others we were a couple
- See less of other friends
- Go out together alone
- Hold hands
- Give partner a present
- Partner gives me a present
- Tell partner love
- Partner tells me love
- Think of ourselves as a couple
- Talk about contraception/STDs
- Would kiss
- Would touch under/no clothing
- Have sex
- Get pregnant
- Get married
Perceptions of Ideal Romance at Wave I, by Emotional Depressive Levels

<table>
<thead>
<tr>
<th>Emotional Depressive Levels</th>
<th>Ideal Romance: What would happen among the first 5 things ideally?</th>
<th>Percent select activity…among the first 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tell partner love</td>
<td>Partner says love me</td>
</tr>
<tr>
<td>BOYS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>24.7</td>
<td>21.5</td>
</tr>
<tr>
<td>Depressed</td>
<td>29.5</td>
<td>24.7</td>
</tr>
<tr>
<td>GIRLS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>11.4</td>
<td>13.0</td>
</tr>
<tr>
<td>Depressed</td>
<td>18.9*</td>
<td>20.0*</td>
</tr>
</tbody>
</table>

Note: * Percentage difference between the depressed and the non-distressed group is significant at < 0.05 level.

### Perceptions of Ideal Romance at Wave I, by Emotional Depressive Levels (Cont.)

<table>
<thead>
<tr>
<th>Emotional Depressive Levels</th>
<th>Ideal Romance: Sexual Contents (%)</th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Having Sex</td>
<td>Touch under/no clothes</td>
<td>Kiss only</td>
<td>Nothing</td>
<td></td>
</tr>
<tr>
<td>BOYS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>47.5</td>
<td>10.9</td>
<td>36.6</td>
<td>5.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>(6,154)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>68.1</td>
<td>8.7</td>
<td>18.8</td>
<td>4.3</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td>(284)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F(2.82, 360.40)=7.8652 p=0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIRLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>29.4</td>
<td>12.7</td>
<td>51.5</td>
<td>6.4</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>(6,435)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>52.6</td>
<td>15.5</td>
<td>29.2</td>
<td>2.7</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>(321)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F(2.84, 363.01)=23.9547 p=0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Actual Romantic Relationships

In-home W1 and W2, R was asked to select and rank activity cards to describe what actually occurred in his/her romantic relationships:

- Went out together in a group
- Met partner’s parents
- Told others we were a couple
- Saw less of other friends
- Went out together alone
- Held hands
- Gave partner a present+
- Partner gave me a present+
- Told partner love*
- Partner told me love*

- Thought of ourselves as a couple
- Talked about contraception/STDs
- Kissed
- Touched under /no clothing
- Touched genitals
- Had sex
- Got pregnant

+ In W2 combined as: Gave each other presents; *Told each other love.
### Actual Romantic Relationships at Wave II by Emotional Depressive Levels at Wave I

<table>
<thead>
<tr>
<th>Emotional Depressive Levels</th>
<th>Actual Romance: What had happened among the first 5 things?</th>
<th>Percent select activity…among the first 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thought of as a couple</td>
<td>Kissed</td>
</tr>
<tr>
<td>BOYS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>51.8</td>
<td>52.2</td>
</tr>
<tr>
<td>Depressed</td>
<td>40.6*</td>
<td>56.2</td>
</tr>
<tr>
<td>GIRLS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Depressed</td>
<td>65.8</td>
<td>61.3</td>
</tr>
<tr>
<td>Depressed</td>
<td>54.5*</td>
<td>64.6</td>
</tr>
</tbody>
</table>

**Notes:** * Percentage difference between the depressed and the non-distressed group is significant at < 0.05 level. + Significance level is at < 0.10.

Actual Romantic Relationships (Cont.)

Other info asked of each partner/relationship:

**Relationship Duration:**
- *Begin date*
- *Still current?*
- *End date*

**Partner demographics:**
- *Current age and grade*
- *Biological sex*
- *Race/ethnicity*

**Sexual Behaviors:**
- *Types of sexual intercourse*
- *1st and most recent dates*
- *Birth control practices*
Actual Romantic Relationships (Cont.)

Other info asked of each partner/relationship (Cont.):

Relationship Context:
- How old was partner?
- What grade?
- Same school? Same church?
- Neighbors? Acquaintances?
- Friends? Friend of friends?
- Close friend(s) knew?

Abuses in Relationship:
- Call names/insult/disrespectful
- Swear at? When 1st happened?
- Threatened w/ violence
- Push/shove
- Throw something at you that hurts?
  When 1st happened?
## Romantic and Sexual Activities in W2 First Reported Relationships, by Race/ethnicity of Couples

**Romance:** Told love, held hands, kissed, told others being a couple, and gave each other gifts  
**Sex:** Touch under/no clothes, genitals, or had sexual intercourse

<table>
<thead>
<tr>
<th>Race/ethnicity of the couple</th>
<th>No Romance, No Sex</th>
<th>No Romance, but Sex</th>
<th>Romance, No Sex</th>
<th>Romance, and Sex</th>
<th>Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both NH-White</td>
<td>22.9</td>
<td>22.4</td>
<td>11.1</td>
<td>43.6</td>
<td>100.0 (3,771)</td>
</tr>
<tr>
<td>Both NH-Black</td>
<td>20.6</td>
<td>32.3****</td>
<td>7.8</td>
<td>39.2</td>
<td>99.9 (1,265)</td>
</tr>
<tr>
<td>Both NH-Asian</td>
<td>32.3</td>
<td>13.9**</td>
<td>14.7</td>
<td>39.1</td>
<td>100.0 (221)</td>
</tr>
<tr>
<td>Both Hispanic</td>
<td>22.9</td>
<td>20.6</td>
<td>20.5***</td>
<td>36.0</td>
<td>100.0 (811)</td>
</tr>
<tr>
<td>Male: NH-White Female: Hispanic</td>
<td>28.6</td>
<td>26.1</td>
<td>9.8</td>
<td>35.5</td>
<td>100.0 (203)</td>
</tr>
<tr>
<td>Male: Hispanic, Female: NH-White</td>
<td>19.6</td>
<td>21.8</td>
<td>10.3</td>
<td>48.3</td>
<td>100.0 (288)</td>
</tr>
</tbody>
</table>

**Notes:** Test of significance by multinominal logistic regression. Contrast group = Couples who are both NH-Whites. Comparisons=odds ratio of being in a specific category (vs. in the base category of 'no romance, no sex.')

**** $p < .001$  
*** $p < .01$  
** $p < .05$
Attitudes towards Marriage/Cohabitation

W3 In-home Data:

- **Important elements for a successful marriage/committed relationship** (Scale 1-10)
  - *Of same race/ethnic group*
  - *Love*
  - *Being faithful*
  - *No cheating*
  - *Make a life-long commitment*
  - *Have enough money*
W3 In-home Data (Cont.):

• **Sense of importance to get married**
  – “I would like to be married now.” (Scale 1-5: Agree/Disagree)
  – “How important is it to you to be married?”
  – “Get married now” vs. “Do other things first”:
    • Finished education
    • Started working full-time for 1-2 years
  – “It is all right for an unmarried couple to live together even if they aren’t interested in considering marriage” (Scale 1-5: Agree/Disagree)
Wave 3 In-home Section 14

Marriage History Begins

H3MR1 How Many Times have you been Married?

0 time
N=12,318

Go to Cohabitation Questions

1, 2, 3 times
N=2,861

Loop through all marriages (up to 3) Marriage A, B, C
Loop through Marriages A, B, C
1st: N=2,861
2nd: N=66
3rd: N=3

H3MR2M_A/B/C
H3MR2Y_A/B/C
Month and Year When Married

H3MR3_A/B/C
Still Married?
1st: N=2,532
2nd: N=49
3rd: N=1

No

H3MR4_A/B/C
How did this marriage end? (Annulment, Divorce, Death)

Yes

H3MR6_A/B/C
Currently living together?

No

H3MR5M_A/B/C
H3MR5Y_A/B/C
Month and Year When Marriage Ended

H3MR7M_A/B/C
H3MR7Y_A/B/C
Month and Year When Last Lived Together
Wave 3 In-home
Section 14
Cohabitation History Begins

H3MR8
Ever Lived w/ someone in marriage-like relationship?
Yes
N=5,993

H3MR9
How Many People?
0 people
N=73

1-10 people
N=5,912

No
N=9,164

Go to Next Set of Questions

Loop through all Cohabitation Episodes A-J (up to 10)

Go to Next Set of Questions
W3 In-home: Data on Cohabitation History (Cont.)

Loop through Cohabitation Episodes A-J

H3MR10M_A to J
H3MR10Y_A to J
Month and Year
When Began
Living Together

H3MR11_A to J
Ever Married?
1st: N=1,020
2nd: N=164
3rd: N=23
...

H3MR12M_A to J
H3MR12Y_A to J
Month and Year
Got Married

IF THIS IS THE
LAST EPISODE/PERSON

NO

H3MR14_A to J
How did this
relationship end?
(R moved out,
partner moved out,
died)

YES

NO

Go to
Next Set of Questions

If YES to H3MR13_A to J
Still living

together?

H3MR15M_A to J
H3MR15Y_A to J
Month and Year
Last Lived together

Frequencies of Cohabitation Episodes:
1st: N=5,910
2nd: N=1,254
3rd: N=252
4th: N=55
5th: N=17
6th: N=7
7th: N=4
8th: N=2
9th: N=2
10th: N=2
W3 In-home Data on Romantic/Sexual Relationships, pregnancies, live births, and children

- Nested Data Structure
  - Romantic/Sexual Relationships since 1995
  - All Pregnancies Associated with Each Relationship/Partner
  - All Live Births (up to 2 for multiple LBs) if Resulted from the Pregnancy
  - All Live Births (not given up for adoption) Who Are Still Living (Children)
W3 In-home Data on Romantic/Sexual Relationships, pregnancies, live births, and children (Cont.)

Partner 1
- Pregnancy 1 → Abortion
- Pregnancy 2 → Multiple Live Births

Partner 2
- Pregnancy 1 → Current Pregnancy

Partner 3
- Pregnancy 1 → Live Birth

Partner 4
- Pregnancy 1 → Live Birth
- Pregnancy 2 → Still Birth

Live Birth 1 → Died
Live Birth 2 → Child 2
Live Birth 1 → Child 1
W3 In-home Data on Romantic/Sexual Relationships, pregnancies, live births, and children (Cont.)

Partner 1
RRELNO=1

- Pregnancy 1
  RPREGNO=1
  - Abortion

- Pregnancy 2
  RPREGNO=2
  - Multiple Live Births

- Current Pregnancy

Live Birth 1
BIRTHNO=1
- Died

Partner 2
RRELNO=2

- Pregnancy 1
  RPREGNO=1
  - Current Pregnancy

Live Birth 2
BIRTHNO=2
- Child 2
  BIRTHNO=2

Partner 3
RRELNO=3

- Pregnancy 1
  RPREGNO=1
  - Live Birth

Live Birth 1
BIRTHNO=1
- Child 1
  BIRTHNO=1

Partner 4
RRELNO=4

- Pregnancy 1
  RPREGNO=1
  - Live Birth

- Pregnancy 2
  RPREGNO=2
  - Still Birth
S17 Relationship Table

- **Unit of Analysis** – Relationship (N=42,334, of which 38,375 were selected for detailed interview in S19).
- **Record ID**: AID + RRELNO
- **Contents**:
  - Relationship current
  - Lasted at least 3 months
  - Male/Female
  - Older/younger/same age
    - by how many years?
  - Race/ethnicity
  - Have had sex
  - Included a pregnancy
    - how many?
W3 In-home: Data on Romantic/Sexual Relationships (Cont.)

S17 Relationship Table (Cont.)

- Ever lived or currently living together
- Ever married or currently married
- Ever adopted a child
- Ever gone to same school between 7-12 grades
- Rank-order partner in having most recent sex
S18 Pregnancy Table

- **Unit of Analysis** – Pregnancy (N=7,422)
- **Record ID**: AID + RRELNO + RPREGNO
- **Contents**:
  - Pregnancy outcomes
    - Miscarriage, abortion, single still birth, live birth, multiple but no live birth, multiple involving both a live birth and another outcome, pregnancy not ended.
    - 7% current; 31% miscarriage/abortion; 58% live births
W3 In-home: Data on Fertility History (Cont.)

S18 Pregnancy Table (Cont.)

• Contents: (Cont.)

  – Pregnancy end or due date (Month/Year)

  – For completed pregnancies with live birth outcome:
    • How many live births per this pregnancy?
      – 4,055 with 1 live birth
      – 173 with multiple live births
W3 In-home: Data on Fertility History (Cont.)

S22 Completed Pregnancies

• **Unit of Analysis** – Pregnancy (N=6,857)

• **Record ID:** AID + RRELNO + RPREGNO

• **Contents:**
  - When s/he was told about pregnancy? [Months into pregnancy]
  - *Female only:*
    - How many weeks when pregnancy ended?
    - Month of pregnancy 1st saw doctor/mid-wife;
    - No of visits.
W3 In-home: Data on Fertility History (Cont.)

S22 Completed Pregnancies (Cont.)

- **Contents: (Cont.)**
  - How far had gone in school when got pregnant?
  - Contacted with partner when pregnancy ended
  - Used birth control then; What kinds;
  - Wanted child then, or later; Wanted partner to be mom/dad of child?
  - R/partner went to prenatal care together; Where to receive prenatal and delivery care; How paid?
  - Drinking, drugs, smoking habits during pregnancy
If pregnancy outcome = live birth, asked:

1. **H3PG22**
   - Were you and partner married at time of this birth?
   - **YES** → Married
   - **NO** →

2. **H3PG23**
   - Were you and partner living together then?
   - **YES** → Cohabitating
   - **NO** →

3. **H3PG24**
   - Which best describes your relationship then?
   - **Did not see/talk to each other**
   - **Hardly saw/talked to each other**
   - **Just friends**
   - **On-again, off-again relationship**
   - **Romantically involved steadily**
### Distribution of Women in Add Health According to their Fertility Experience and Relationship Context of First Birth

<table>
<thead>
<tr>
<th>First Births by Relationship Context</th>
<th>%</th>
<th>% Among Women with a First Birth</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marriage</strong></td>
<td>7.3</td>
<td>28.0</td>
<td>409</td>
</tr>
<tr>
<td><strong>Cohabitation</strong></td>
<td>8.1</td>
<td>30.4</td>
<td>444</td>
</tr>
<tr>
<td><strong>Dating</strong></td>
<td>4.1</td>
<td>18.2</td>
<td>266</td>
</tr>
<tr>
<td><strong>No Relationship</strong></td>
<td>5.7</td>
<td>23.3</td>
<td>340</td>
</tr>
<tr>
<td><strong>Never had a live birth</strong></td>
<td>74.8</td>
<td></td>
<td>4,429</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100.0</td>
<td>100.0</td>
<td>5,888</td>
</tr>
</tbody>
</table>

Relationship between Family of Origin Family Structure and Women’s Fertility Experience and Relationship Context of First Birth

W3 In-home: Data on Fertility History (Cont.)

S23 Current Pregnancies

• **Unit of Analysis** – Pregnancy (N=534)
• **Record ID**: AID + RRELNO + RPREGNO
• **Contents:**
  – Told partner about pregnancy; When [Months into pregnancy]?
  – What month of pregnancy now?
  – Seen doctor/mid-wife yet; Any plan?
  – No. of visits so far; Gone with partner; Where to receive prenatal and delivery care; How paid?
  – Drinking, drugs, smoking habits during pregnancy
S23 Current Pregnancies (Cont.)

• Contents: (Cont.)
  – How far has gone in school?
  – How old when got pregnant?
  – Used birth control then; What kinds; Wanted child then, or later; Want partner to be mom/dad of child?
  – Relationship context:
    • Married
    • Cohabiting
    • Never/hardly saw/talked
    • Just friends
    • On-and-off again
    • Romantically involved steadily
W3 In-home: Data on Fertility History (Cont.)

S24 Live Births

- **Unit of Analysis** – Live birth (N=4,336)
- **Record ID:** AID + RRELNO + RPREGNO + BIRTHNO
- **Contents:**
  - Baby went home; Why not?
  - Boy/Girl; R named baby
  - Birth weight and length
  - Born early; How many weeks?
  - Any major health problems; Still living; When died?
  - *[Male]* Partner present at birth
    - Name on birth certificate
    - Baby took on his last name
S25 Children and Parenting

- **Unit of Analysis** – Child (N=4,181)
- **Record ID**: AID + RRELNO + RPREGNO + BIRTHNO
- **Contents**:
  - How old?
  - Child lives with R; Ever?
    - Last date living together
    - Distance
    - With whom child lives with
  - R lives with partner? Ever?
    - Last date living together
    - Distance
S25 Children and Parenting (Cont.)

• Contents: (Cont.)
  
  – Legal agreement with partner
    • Spent time with child
    • Last seeing child
    • Child support
  
  – Child’s health; Disability; Limitation
  
  – A list of health problems
  
  – Caring/parenting problems
  
  – Language spoken to child
Computational and Technical Tips

Data Structure Transformation for Hierarchical Files

1. Construct unique record IDs for each file

<table>
<thead>
<tr>
<th>Type of ID</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual ID</td>
<td>Individual ID (AID)</td>
</tr>
<tr>
<td>Relationship ID</td>
<td>AID + Relationship ID (AID + RRELNO)</td>
</tr>
<tr>
<td>Pregnancy ID</td>
<td>AID + Relationship ID + Pregnancy ID (AID + RRELNO + RPREGNO)</td>
</tr>
<tr>
<td>Live birth ID</td>
<td>AID + Relationship ID + Pregnancy ID + Live birth ID (AID + RRELNO + RPREGNO + BIRTHNO)</td>
</tr>
<tr>
<td>Children ID</td>
<td>AID + Relationship ID + Pregnancy ID + Live birth ID + Children ID (AID + RRELNO + RPREGNO + BIRTHNO)</td>
</tr>
</tbody>
</table>
2. Lower level file contains IDs of all higher level files

For example:

Pregnancy records contain record IDs of the corresponding individual, relationship, besides its own unique pregnancy record ID.
Data Structure Transformation for Hierarchical Files (Cont.)

3. Example: Construct a pregnancy history for R

   a. Chronologically rank-order pregnancies associated with R by the pregnancy end date
   b. Decide what pregnancy variables to write out
      – Pregnancy end date; pregnancy outcome
      – Retain Pregnancy ID (for linking back to the original pregnancy data, or to the associated relationship, or live birth records, if ever needed)
   c. Transform data from “long” to “wide”
Computational and Technical Tips (Cont.)

Data Structure Transformation for Hierarchical Files (Cont.)

4. Example: Construct a pregnancy history for R (Cont.)

Using SAS codes

a. Construct pregnancy record ID (a character variable) by putting together AID, RRELNO, RPREGNO

b. Compute SAS dates for pregnancy end date

c. Use PROC SORT to chronologically rank-order pregnancies associated with each R.

d. Transform data from “long” to “wide” using ARRAY, RETAIN, and OUTPUT statements.
Compute Summary Counts

1. Describe R by his/her reporting on total numbers of ‘event’: For examples
   a. Total number of relationships reported by R
   b. Total number of pregnancies reported by R

2. Describe R by his/her average reporting on a certain ‘event’: For example
   a. Average number of pregnancy per relationship by R
3. Example: Compute total no. of relationship per R

Using SAS codes
a. PROC SORT Relationship records by AID
b. Use SUM command in PROC SUMMARY to compute and OUTPUT total no. of records per R.
c. Merge output summary variable to individual level file by AID
Compute Summary Counts (Cont.)

4. Example: Compute total no. of pregnancy per R

Using SAS codes

a. PROC SORT Pregnancy records by AID

b. Use SUM command in PROC SUMMARY to compute and OUTPUT total no. of records per R.

c. Merge output summary variable to individual level file by AID
Compute Summary Counts (Cont.)

5. Example: Compute average no. of pregnancy per relationship by R

Using SAS codes

a. Use the individual level file with the two merged summary variables (total no. of pregnancies per R, total no. of relationships per R).

b. Compute average no. of pregnancy per relationship by dividing total no. of pregnancy by total no. of relationship.
Some Data Cleaning and Programming Tips

1. Skipped-pattern checks
   a. Beware of the DK/RF/NA and odd missing cases.
   b. Know your denominators: Based on what did you take the percentage of?
   c. Check the branching of your variables: Who are the legitimate cases?

Using SAS codes
   a. Use LIST MISSPRINT commands in PROC FREQ to trace all legitimate cases for each variable.
2. Range checks
   a. Not only basic range check at univariate level, but also cross-classified by relevant variables for reasonable range check (data inconsistency).
   b. Investigate and make decisions on extreme/odd cases

Using SAS codes
   a. Use IF, THEN, PROC PRINT statements to list off all out-of-range or extreme/odd cases.
   b. Use FORMAT statements to facilitate easier reading of a long list of values during case-by-case investigation.
3. Check duplicates
   a. Why duplicate records possible? Special problems for hierarchical files with records automatically generated by R’s answers to prior questions.
   b. Duplication cannot be identified simply by unique record IDs.
   c. Find possible duplications by comparing response patterns across records.
Some Data Cleaning and Programming Tips (Cont.)

3. Check duplicates: (Cont.)
   
d. Once isolate ‘bad records’, check against all other related data records (across hierarchical levels) associating w/ the same R.

e. Define what constitutes a probable duplication and apply general rules consistently to all cases.

f. Flag records with different degrees of data problems.

g. Remove duplicates before data analysis.
Some Data Cleaning and Programming Tips (Cont.)

4. An example: Check duplicate pregnancy records
   a. Compare pregnancy end/due date with interview date. Resolve data inconsistency (e.g., completed pregnancy ended after interview or current pregnancy due prior to interview).
   b. Chronologically rank-order pregnancies.
   c. Compare pregnancy date across records and flag those with ‘too close’ a duration in-between pregnancies. Flag the ‘too close’ records.
4. An example: Check duplicate pregnancy records (Cont.)

‘Too Close’ a time gap as defined by –

<table>
<thead>
<tr>
<th>Regardless ‘prior’ pregnancy outcome, minimum time gap allowed for before the ‘next’ pregnancy end date:</th>
<th>Pregnancy outcome of the ‘next’ pregnancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 months</td>
<td>Abortion, miscarriage</td>
</tr>
<tr>
<td>10 months</td>
<td>Current pregnancy, live birth, multiple with at least one live birth</td>
</tr>
<tr>
<td>7 months</td>
<td>Single, still birth</td>
</tr>
<tr>
<td>7 months</td>
<td>Multiple, no live birth</td>
</tr>
</tbody>
</table>
4. An example: Check duplicate pregnancy records (Cont.)

d. Compare response patterns item-by-item.

e. Compute no. of same responses across records.

f. Examine total no. of DK, RF, NA responses in each record.

g. Flag records with ‘too close’ a time gap between pregnancies, 90% more similar response patterns, and 80% more DK/RF/NA responses.
Computational and Technical Tips (Cont.)

Some Data Cleaning and Programming Tips (Cont.)

4. An example: Check duplicate pregnancy records (Cont.)
   h. List off these cases with their corresponding relationships, live births, and child records for further examination.
   i. Set up cleaning rules for final deletion of ‘complete/partial’ duplicates.
Add Health Wave IV

- Data collection 2007-08 when cohort members are 24-32 and completing transition to adulthood.
- Follow all Wave I respondents.
- Research Triangle Institute (RTI) is the Wave IV field contractor.
Wave IV Questionnaire Content

- Overview & demographics
- Household roster & residence history
- Relationships w/ parents and siblings
- Education
- Labor market, military experience
- Economics
- General health & diet
- Access to health services
- Illnesses & medications
- Personality

- Stressors & mental health
- Cognitive function
- Relationship, sexual, & fertility histories
- Children & parenting
- Delinquency & violence
- Involvement w/ criminal justice system
- Activity & sleep
- Substance use & abuse
- Religiosity & civic participation
- Future expectations
Wave IV: A Preview on the 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. person(s) romantically/sexually involved since 2001

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

M partners

No. of person(s) ever married

C partners

No. of person(s) ever lived with

P partners

No. of person(s) resulted in a pregnancy

CD partners

No. of person(s) in current relationships

MR partner

No. person(s) romantically/sexually involved since 2001
Wave IV: A Preview on the Hierarchical Data Structure of Individual, Relationship, Pregnancy, Live Birth, and Child Records

- 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
- 1st time
- 2nd time
- Nth time...
- Begin/end Dates, Duration
- Begin/end Dates, Duration
- Begin/end Dates, Duration
- Begin/end Dates, Duration
- Begin/end Dates, Duration
- Begin/end Dates, Duration
- Begin/end Dates, Duration
- Begin/end Dates, Duration
Wave IV: A Preview on the Hierarchical Data Structure of Individual, Relationship, Pregnancy, Live Birth, and Child Records

- M partners: How many times got/made pregnant by partner?
- C partners: How many times got/made pregnant by partner?
- P partners: How many times got/made pregnant by partner?
- CD partners
- MR partner

Pregnancy Records

Loop through M, C, P partners only
Wave IV: A Preview on the Hierarchical Data Structure of Individual, Relationship, Pregnancy, Live Birth, and Child 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

**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
Wave IV: A Preview on the Hierarchical Data Structure of Individual, Relationship, Pregnancy, Live Birth, and Child Records

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

If sum of alive BB (for all live births) per R = 0 -> true

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

Loop through all child records per R

At Individual Level, ask:
Language use at home
Parenting experiences

Skip Child & Parenting Section
If = 0

Check Household Roster for any Step/adopted children reported
Wave IV Data
Expected Release: Fall 2009

- Longitudinal survey data
- Geographic data
- Biological data
Some General Advices

• Do not assume: Read as much as you can

  – Add Health documentation on website information, documents, codebook).

  – When in doubt, ask Add Health data manager.

  – Subscribe to Add Health list server.
Some General Advice

• Spend time: To know your data

  – Examine question branching; Test skipped 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!
Some General Advice (Cont.)

• 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.

  – Organize your programs and results, perhaps using a electronically searchable research management database.
Some General Advice (Cont.)

• Join the excitement and fun, and be in the Add Health User Group!
  – Check out publications using Add Health data.
  – Be on the look out for the Wave IV data release.
  – Attend Add Health Users Conferences in Bethesda, MD.
Questions and Answers