

Strategies for Constructing Household and Family Units with Linked Administrative Records

AISP Presentation

Regan Foust

Stephanie Cuccaro-Alamin

January 25th, 2021



Harnessing the scientific potential of linked, administrative data to inform children's programs and policies.

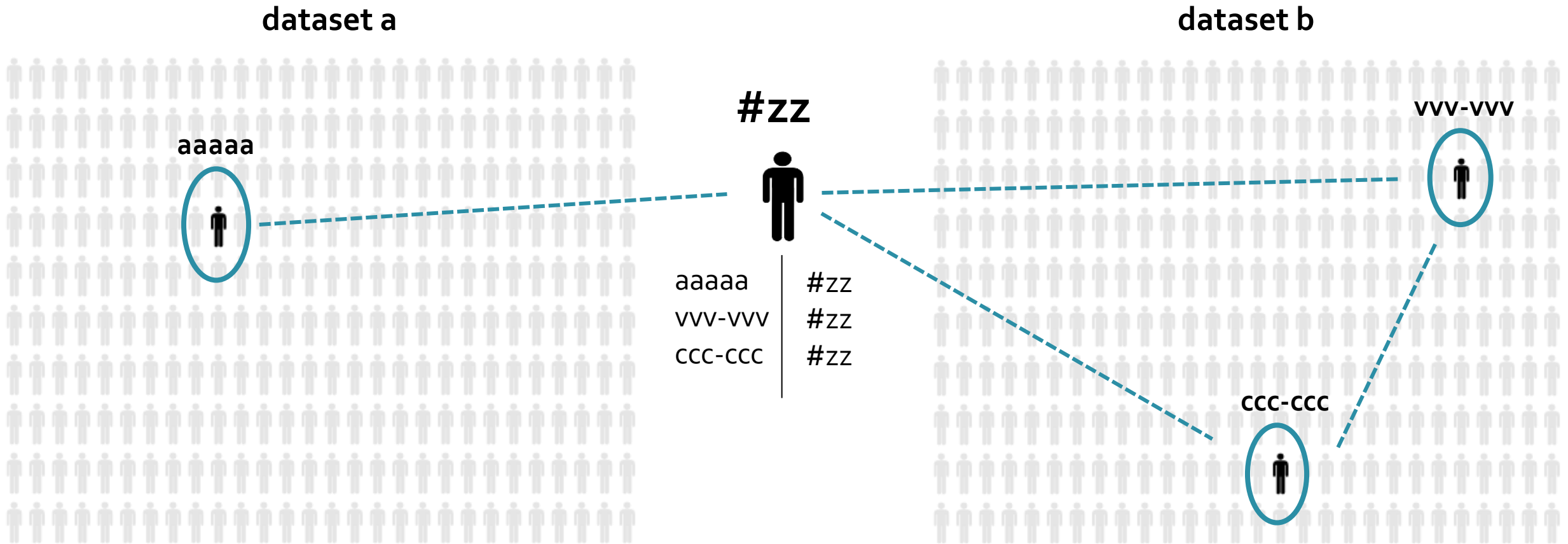




CHHS Record Reconciliation Project



probabilistic de-duplication and cross-program linkage of records using an algorithm trained / customized for California data (*machine learning*)



2018

**1st record
reconciliation**

2016 records

Medi-Cal records

CalFresh records

Child Welfare records

Developmental Service records

IHSS records

CalWORKs records

FPACT records

WIC records

2019

**2nd record
reconciliation**

[+] 2015, 2017, 2018 records

[+] **vital birth** records

[+] **vital death** records

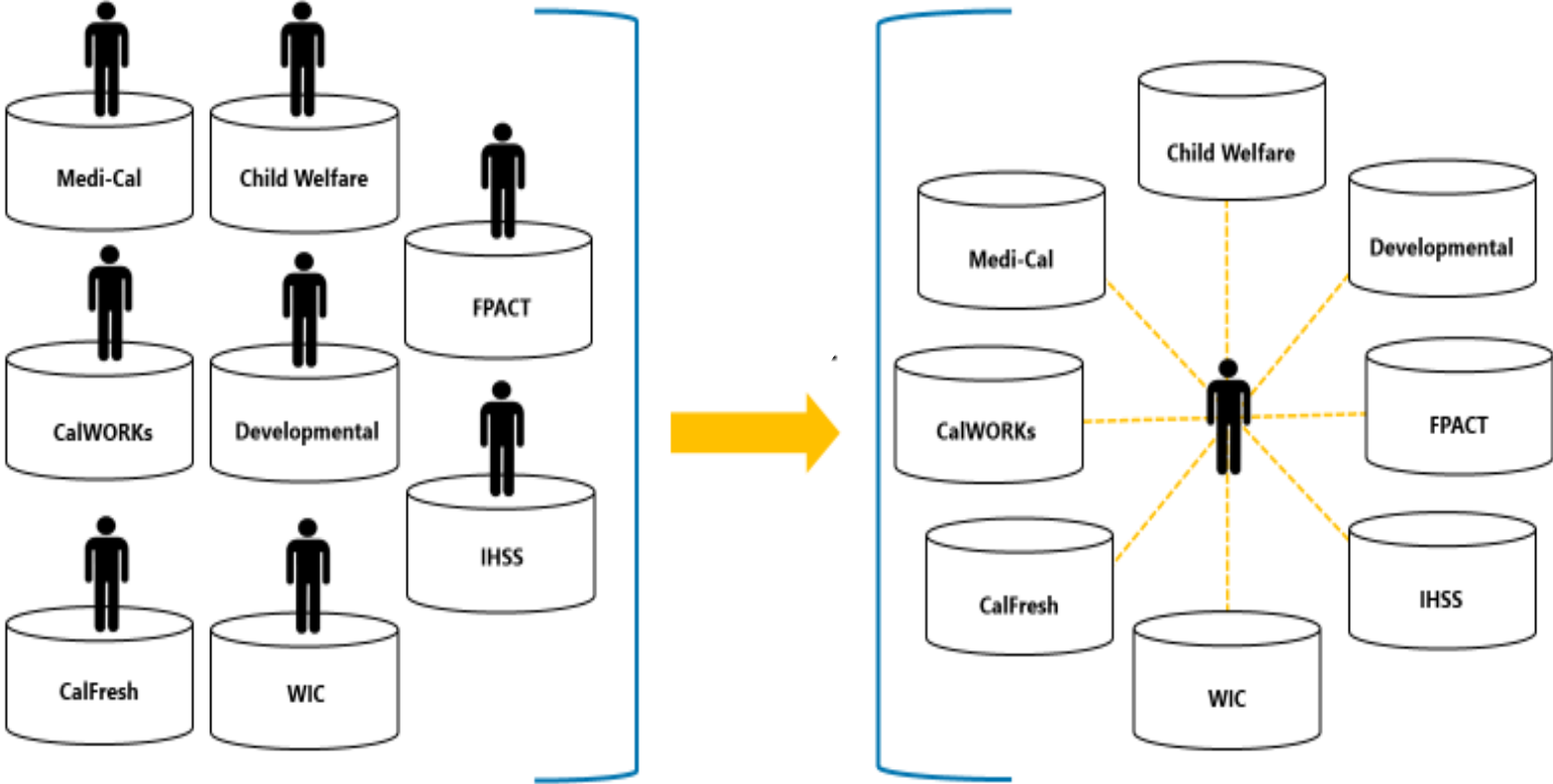
2020

**3rd record
reconciliation**

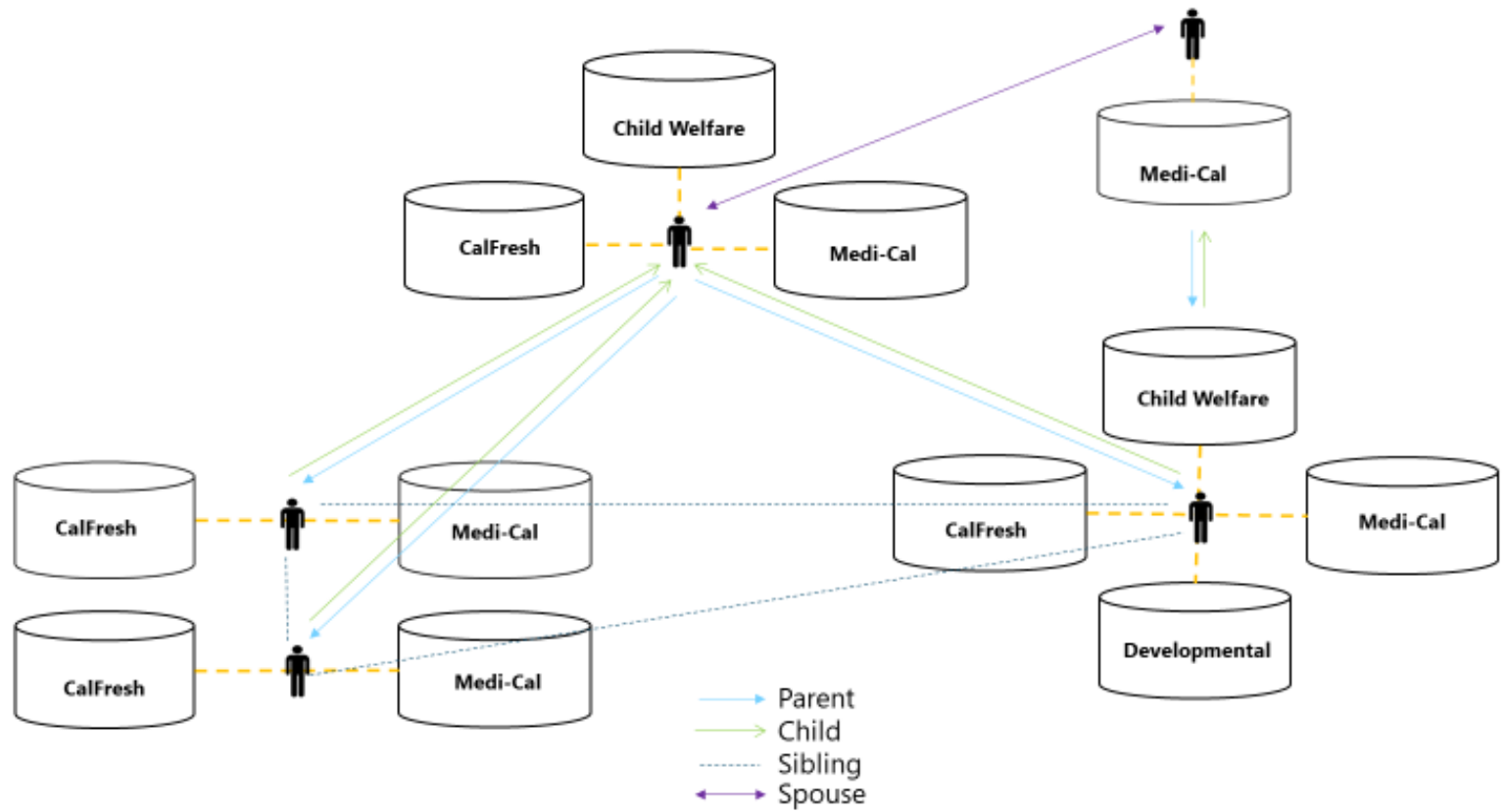
[+] 2019 records

Data Structure

Data are inherently “program-centric” but the record reconciliation effort makes data “client centered”



Limitation

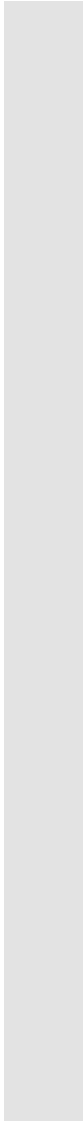


A client-centered structure doesn't shed light on the important relationships *among* the individuals in these data.



Research Review

The objective was to outline conceptual strategies that could be adopted for organizing CHHS records into households and/or family units.



Purpose

- Provide an understanding of **how households and family units are conceptualized** in the United States;
- Illustrate the **limitations** of this approach for measuring **complex families**;
- Explore **strategies** for organizing administrative data into household and family units for this purpose; and
- Determine which of these strategies might be **applicable to CHHS data**.

Measuring Family

- **Family** is the basic unit of social and economic organization in society and the primary mechanism for human socialization, support, and development.
- How families are **operationalized in research** – specifically what boundaries are drawn around them for measurement – is critical.
- Determining “**Who Counts?**” has important implications for our understanding of family functioning.
- But defining family can be problematic.
 - Family is **subjective** (It depends on **who** (?) you ask)
 - Family is **dynamic** (It depends on **when** (?) you ask)

Census Approach

- Very **influential** in shaping the way we define and measure families.
- Rooted in neo-classical economics and revolves around the **household or residence**.

Please do NOT photocopy this questionnaire. Each questionnaire has a unique ID number. OMB No. 2607-1026 Approval Expires 11/2022

United States
Census 2020

U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU

Individual Census Questionnaire

Start here Use a blue or black pen.

NOTE: Please answer BOTH Question 6 about Hispanic origin and Question 7 about race. For this census, Hispanic origins are not races.

1. What is your name? Print name below.

First Name MI LL SS JR SR III IV V VI VII VIII IX X XI XII XIII XIV XV XVI XVII XVIII XIX XX XXI XXII XXIII XXIV XXV XXVI XXVII XXVIII XXIX XXX XXXI XXXII XXXIII XXXIV XXXV XXXVI XXXVII XXXVIII XXXIX XL XLI XLII XLIII XLIV XLV XLVI XLVII XLVIII XLIX L LI LII LIII LIV LV LVI LVII LVIII LIX LX LXI LXII LXIII LXIV LXV LXVI LXVII LXVIII LXIX LXX LXXI LXXII LXXIII LXXIV LXXV LXXVI LXXVII LXXVIII LXXIX LXXX LXXXI LXXXII LXXXIII LXXXIV LXXXV LXXXVI LXXXVII LXXXVIII LXXXIX XLXXX XLXXXI XLXXXII XLXXXIII XLXXXIV XLXXXV XLXXXVI XLXXXVII XLXXXVIII XLXXXIX LXXXX LXXXXI LXXXXII LXXXXIII LXXXXIV LXXXXV LXXXXVI LXXXXVII LXXXXVIII LXXXXIX LXXXXX LXXXXXI LXXXXXII LXXXXXIII LXXXXXIV LXXXXXV LXXXXXVI LXXXXXVII LXXXXXVIII LXXXXXIX LXXXXXX LXXXXXXI LXXXXXXII LXXXXXXIII LXXXXXXIV LXXXXXXV LXXXXXXVI LXXXXXXVII LXXXXXXVIII LXXXXXXIX LXXXXXXX LXXXXXXXI LXXXXXXXII LXXXXXXXIII LXXXXXXXIV LXXXXXXXV LXXXXXXXVI LXXXXXXXVII LXXXXXXXVIII LXXXXXXXIX LXXXXXXX

Last Name(s) _____

2. Do you live or stay here **most of the time**?
 Yes No

3. Besides here, what is the full address of a place where you sometimes live or stay?
 I never stay at any other place. I only live here.

Address Number (For example: 3007) _____

Street Name (For example: N Maple Ave) _____

Apt./Unit (For example: Apt A or Unit 2) _____

Rural Route Address (if there is no street address) _____

City _____

State _____ ZIP Code _____

NOTE: Please provide a location description below if there is no address or if this is a facility.

4. What is your sex? Mark **ONE** box.
 Male Female

5. What is your age on April 1, 2020, and what is your date of birth? If you don't know the exact age, please estimate. For babies less than 1 year old, do not write the age in months. Write 0 for the age.

Age on April 1, 2020 _____

Print numbers in boxes.
Month _____ Day _____ Year of birth _____

6. Are you of Hispanic, Latino, or Spanish origin?
 No, not of Hispanic, Latino, or Spanish origin
 Yes, Mexican, Mexican Am., Chicano
 Yes, Puerto Rican
 Yes, Cuban
 Yes, another Hispanic, Latino, or Spanish origin -- Print, for example, Salvadoran, Dominican, Colombian, Guatemalan, Spanish, Ecuadorian, etc. _____

7. What is your race? Mark **X** one or more boxes **AND** print origin.
 White -- Print, for example, German, Irish, English, Italian, Latvian, Egyptian, etc. _____
 Black or African Am. -- Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc. _____
 American Indian or Alaska Native -- Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec; Native Village of Barrow Inupiat Traditional Government, Nome Salween Community, etc. _____
 Chinese Vietnamese Native Hawaiian
 Filipino Korean Samoan
 Asian Indian Japanese Chamorro
 Other Asian -- Print, for example, Pakistani, Cambodian, Hmong, etc. _____
 Other Pacific Islander -- Print, for example, Tongan, Fijian, Marshallese, etc. _____
 Some other race -- Print race or origin _____

DI-Q-GE 01-20-2020

U.S. Census Household and Family Definitions

Household – “a group of people who share a **residence** in common and such things as meals, bills, facilities or chores, but it also could describe one person living alone”.

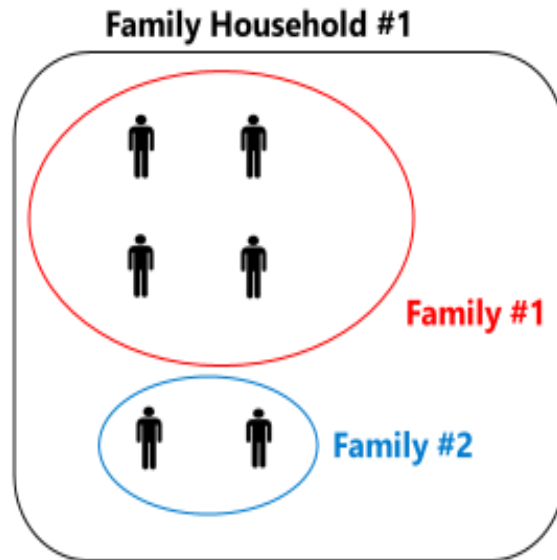
- family household
- non-family household

Family – “A householder and one or more people living in the same household who are related to the householder.” All people in a household who are **related** to the householder (by birth, marriage, or adoption) are regarded as members of his or her family.

Differences between Households and Families

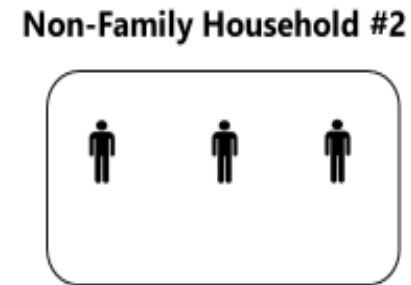
- The term **family** is often used interchangeably with the word **household**; and although both are both basic units of analysis in demography, they are not the same thing.
- Understanding how these two concepts overlap and diverge is critical to understanding how families are measured.
- In short, **households can contain multiple families - and families can exist across multiple households.**

Traditional Methods for Defining Households and Families

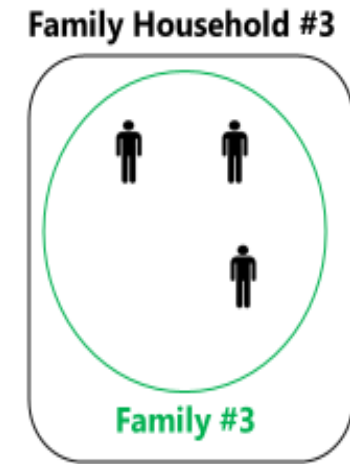


Family Household #1 – 6 members

Family #1 – 4 members
Family #2 – 2 members



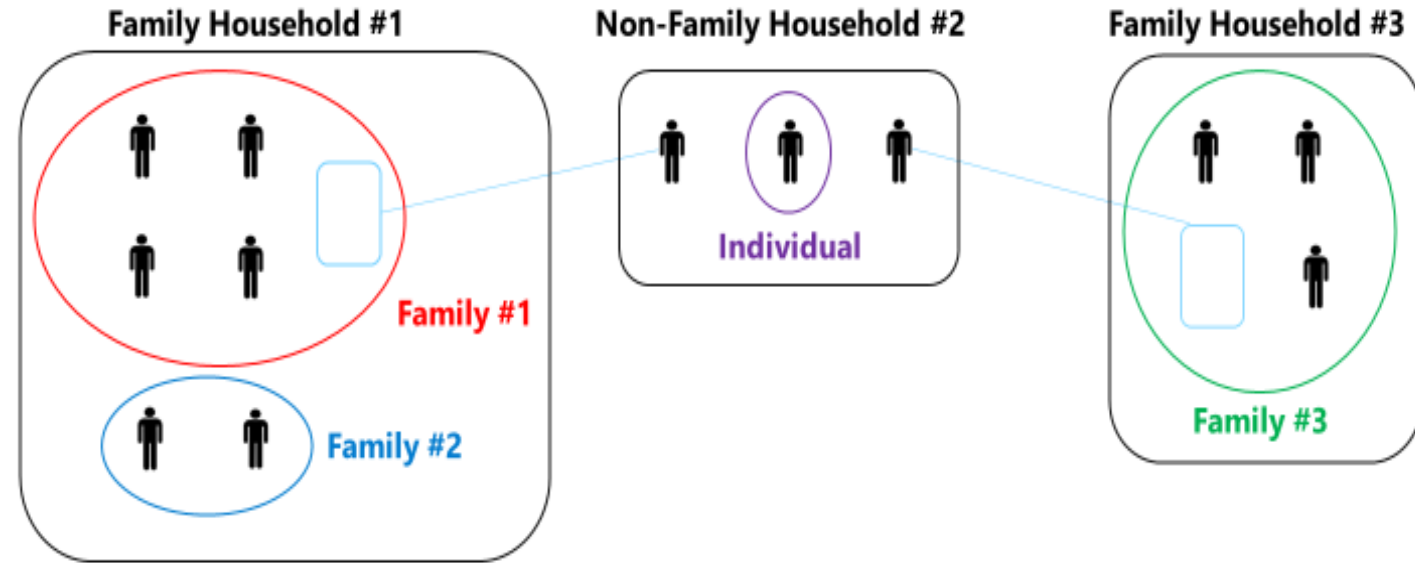
Non-Family Household #2 – 3 members



Family Household #3 – 3 members

Family #3 – 3 members

Households and Families in Reality



Family Household #1 – 6 members

Non-Family Household #2 – 3 members

Family Household #3 – 3 members

Family #1 – 5 members, 2 households

Family #2 – 2 members, 1 household

Family #3 – 4 members, 2 households

Family Structure

- **Family structure** refers to the “composition and membership of the family and the organization and patterning of relationships among individual family members.”
- Census Approach - classifications of family structure focus on:
 - number of parents
 - marital/union status of parents
 - relationship of parents to children
- Classifications are not always consistent with **people’s experience of family**.

Characteristics of Families

Dynamic

- Families are geographically dynamic - they may reside across several households, even states or continents.
- Families are also temporally dynamic - they experience transitions such as births, changes in marital status, or living arrangements (e.g., cohabitation).

Ambiguous boundaries

- Boundaries within families are often more "flexible and permeable" than rigid household constructs can measure.
- When defined objectively, family membership is always reciprocal, but in reality, family relationships are asymmetrical.
- **Demographic shifts** over the last half-century have produced increasingly **complex family structures**.
- **Census approach is not always able to capture these complex features.**

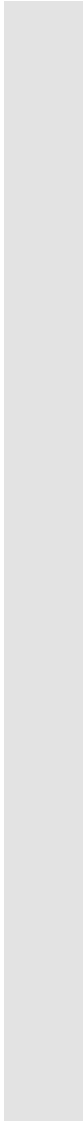
Other Approaches

- Researchers have worked to develop **alternate approaches** to characterize more complex family forms. But these methods are **difficult to apply** to large scale data systems.
- Increasingly researchers are turning to **linked administrative data** for research. Advantages include broad coverage, inexpensive, reliable, longitudinal and typically cover hard to reach populations of interest.
- Improvements in data collection and increasingly sophisticated linkage techniques have made **linked administrative data** an **increasingly promising method** for household and family research.



Strategies

Organizing administrative data into household and family units for analysis using linked administrative data



Overview

- We identified **two preferred strategies** for constructing household and family units that could be implemented using information available in **CHHS records**.
- **Neither are 'perfect'** in their ability to adequately capture the intricacies of complex families, but each has its own particular advantages, and could, potentially, be **employed in tandem** to answer questions of interest.

Strategies

Strategy 1 Household Method

- uses shared address to define Households
- assumes that families are synonymous with households
- household composition / family structure is defined within the shared residence only

Strategy 2 Family Relationship Method

- uses family relationships to define Families
- prioritizes family (e.g., biological, adoptive, or foster) relationships between individual clients in the CHHS data
- family structure is defined within a family roster

Strategy 1

**Household
Method**



Step 1

Create Person File

Construct a file of every unique individual available across all program data sets.

- Split the birth record into three separate records. One each for child, mother (or first parent), and father (or second parent) to obtain a universe of individuals defined as a family unit at birth.
- Once linkages have been completed, the complete universe of all unique clients in the CHHS record reconciliation is merged in to create the person file.

Figure	DOB	DOD	Gender
Joan	1/11/1983		F
Mike	8/22/1978		M
Patty	10/22/2010		F

Step 2

Create Address File

Construct a file that includes every unique address available across all program data sets.

- Addresses can be geocoded to augment the file with census tract information and other geographic data.

Address	ADD ID	Census Tract	Coordinates
222 X Street NW, Anytown, CA 94111	ADD1	48-201-533701	TIGER Info
222 X Street, Anytown, 94111	ADD1	48-201-533701	TIGER Info
35 T Ave, Anytown, CA 94122	ADD2	48-201-533567	TIGER Info

Step 3

Create Individual Household File

Construct a file to track each individual's residence history.

- Using the person file as the initial spine, each unique individual is linked to each of the CHHS program source files and an event level record is created for every occurrence of an address recorded for that individual in that specific service program.
- Using available service dates, approximate address episode start and end dates fields would be created for each record to indicate the first and last recorded date for the individual at a specific address.
- As new data sources are integrated into the CHHS record reconciliation, information regarding an individuals address history can be improved.

Client ID	DOB	DOD	Gender	Address	ADD ID	Case ID	P start date	P end date	E start date	E end date
Joan	1/11/1983		F	222 X Street NW, Anytown, CA 94111	ADD1		10/22/2010		10/22/2010	3/1/2017
Joan	1/11/1983		F	222 X Street NW, Anytown, CA 94111	ADD1		1/1/2016	9/19/2016	5/25/2014	3/1/2017
Joan	1/11/1983		F	222 X Street, Anytown, 94111	ADD1	HYJK56R	9/19/2016	3/1/2017	5/25/2014	3/1/2017
Mike	8/22/1978		M	222 X Street NW, Anytown, CA 94111	ADD1		10/22/2010		10/22/2010	3/1/2017
Mike	8/22/1978		M	222 X Street NW, Anytown, CA 94111	ADD1		1/1/2016	9/19/2016	5/25/2014	3/1/2017
Mike	8/22/1978		M	222 X Street, Anytown, 94111	ADD1	HYJK56R	9/19/2016	3/1/2017	5/25/2014	3/1/2017
Patty	10/22/2010		F	222 X Street NW, Anytown, CA 94111	ADD1		10/22/2010		10/22/2010	11/19/2016
Patty	10/22/2010		F	222 X Street NW, Anytown, CA 94111	ADD1		1/1/2016	11/19/2016	5/25/2014	11/19/2016
Patty	10/22/2010		F	222 X Street, Anytown, 94111	ADD1	HYJK56R	9/19/2016	11/19/2016	5/25/2014	11/19/2016
Patty	10/22/2010		F	35 T Ave, Anytown, CA 94122	ADD2	89ujjy65	11/19/2016	1/5/2017	11/19/2016	1/5/2017
Patty	10/22/2010		F	35 T Ave, Anytown, CA 94122	ADD2	56723chs	11/19/2016	1/15/2017	11/19/2016	1/15/2017

Step 4

Group Households

Household is defined as “a group of individuals who share the same unique address at a particular point in time.

- Records from the individual household file are first sorted by address ID, CHHS Client ID, and program start dates and end dates.
- Records where individuals are known to be residing at an address on a specific date or during the specified time period are selected.

Client ID	DOB	DOD	Gender	Address	ADD ID	Case ID	P start date	P end date	E start date	E end date
Joan	1/11/1983		F	222 X Street NW, Anytown, CA 94111	ADD1		10/22/2010		10/22/2010	3/1/2017
Joan	1/11/1983		F	222 X Street NW, Anytown, CA 94111	ADD1		1/1/2016	9/19/2016	10/22/2010	3/1/2017
Joan	1/11/1983		F	222 X Street, Anytown, 94111	ADD1	HYJK56R	9/19/2016	3/1/2017	10/22/2010	3/1/2017
Mike	8/22/1978		M	222 X Street NW, Anytown, CA 94111	ADD1		10/22/2010		10/22/2010	3/1/2017
Mike	8/22/1978		M	222 X Street NW, Anytown, CA 94111	ADD1		1/1/2016	9/19/2016	10/22/2010	3/1/2017
Mike	8/22/1978		M	222 X Street, Anytown, 94111	ADD1	HYJK56R	9/19/2016	3/1/2017	10/22/2010	3/1/2017
Patty	10/22/2010		F	222 X Street NW, Anytown, CA 94111	ADD1		10/22/2010		10/22/2010	11/19/2016
Patty	10/22/2010		F	222 X Street NW, Anytown, CA 94111	ADD1		1/1/2016	11/19/2016	10/22/2010	11/19/2016
Patty	10/22/2010		F	222 X Street, Anytown, 94111	ADD1	HYJK56R	9/19/2016	11/19/2016	10/22/2010	11/19/2016
Patty	10/22/2010		F	35 T Ave, Anytown, CA 94122	ADD2	89ujjy65	11/19/2016	1/5/2017	11/19/2016	1/5/2017
Patty	10/22/2010		F	35 T Ave, Anytown, CA 94122	ADD2	56723chs	11/19/2016	1/15/2017	11/19/2016	1/15/2017

Step 5

Define Household Roster

- Records are then assigned a **Household ID** that combines the address identifier and date information.
- Records can then be unduplicated by selecting one record per client within the household.
- This table represents the Household Roster as of July 1, 2016

Client ID	DOB	DOD	Gender	ADD ID	ADD_house
Joan	1/11/1983		F	ADD1	ADD1_O70116
Mike	8/22/1978		M	ADD1	ADD1_O70116
Patty	10/22/2010		F	ADD1	ADD1_O70116

Step 6

Assign Household Types

- Once records are unduplicated at the household level, summary measures of **household composition** can be defined.
 - total number of clients in the household (size)
 - number of members age 0-17, 18-64 and 65+.
- Following the approach similar to Harper and Mahew (2016), this summary information can then be configured to define **core household types (A-I)**.

Household Type	Size	age 0-17	age 18-64	age 65+
Couple with (4) children (A.4)	6	XXXX	XX	
Couple with (3) children (A.3)	5	XXXX	XX	
Couple with (2) children (A.2)	4	XX	XX	
Couple with (1) children (A.1)	3	X	XX	
Single adult with one child (B)	2	X	X	
Older couple with 1 person aged 65+ (C)	2		X	X
Older person living alone (D)	1			X
3-gen with one child, couple and an older person (E)	4	X	XX	X
Cohabiting adult (F)	3		XXX	
Adult living alone (G)	1			X
Split generation (H)	4	XX		XX
Young (e.g., students, teenage parents) (I)	4	XXXX		

Source: Adapted from Harper & Mayhew (2016) p. 6

Advantages

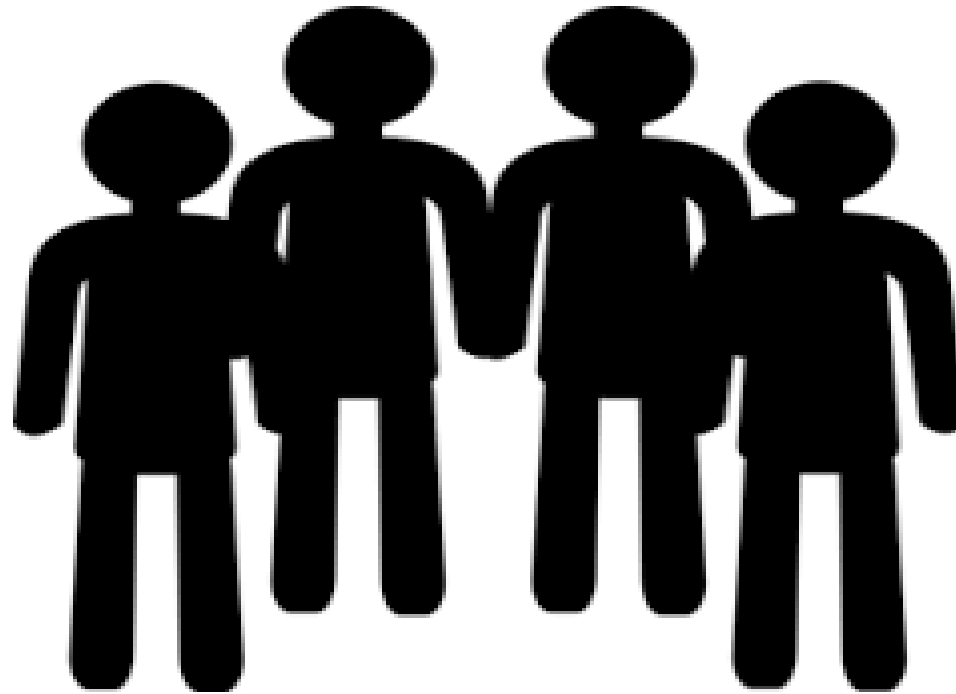
- Using the Unique Household ID, this information can then be **linked back to the Client's individual CHHS Linkage Key** so that information for the household members from the component service programs can be examined.
- Because it is not based on biological relationship alone, the approach can **account for non-relatives and extended family members** who might live in the household.
- Because individuals can be recorded in more than one household / address at a specific point in time, the prevalence of **complex living arrangements** can also be assessed.

Limitations

- Because the approach conflates household with family, it may **miss important family members who reside in other households**, such as children who are away at school or been placed in out-of-home care.
- Does not provide information regarding the **relationships between household members**. Therefore, doubled up families who inhabit the same household may be classified as one, or inaccurate address data for multi-family units might result in residents of different units being classified as one household.
- Because dates of service are used to place individuals in households, **missing or incorrect dates** might result in two different “families” being grouped in the same household if they reside at one address in succession.
- Selection of **longer analysis time periods** could also result in inaccurate classification of households as dates of residence might overlap.

Strategy 2

Family
Relationship
Method



Step 1

Create Person File

Construct a file of every unique individual available across all program data sets.

This is the same as the first step of the Household Strategy (#1)

Client ID	DOB	DOD	Gender
Sue	1/11/1983		F
Joe	8/22/1978		M
Shelly	5/25/2014		F

Step 2

Create Client Relationship File - Birth

Construct a file to record each individual's birth relationships.

- For each primary client in the person file, the individual identified as mother or father will be recorded as the secondary client and a relationship field will be coded bio daughter or bio son depending on the gender of the primary client.
- Relationship is always defined **directionally** from the perspective of the primary client (i.e., "what relationship is the secondary client to the primary client?").

Primary Client	Primary DOB	Primary Gender	Secondary Client	Secondary DOB	Secondary Gender	S:P Relation	Start	End
Shelly	5/25/2014	F	Joe	1/11/1983	F	bio mother	5/25/2014	
Shelly	5/25/2014	F	Joe	8/22/1978	M	bio father	5/25/2014	
Sue	1/11/1983	F	Shelly	5/25/2014	F	bio daughter	5/25/2014	
Joe	8/22/1978	M	Shelly	5/25/2014	F	bio daughter	5/25/2014	

Step 3

Add Sibling Relationships

- The file is sorted by secondary client ID in order to **group siblings who share a biological parent.**

Primary Client	Primary DOB	Primary Gender	Secondary Client	Secondary DOB	Secondary Gender	S:P Relation	Start	End
Meg	7/19/2015	F	Frank	11/19/2011	M	bio father	7/19/2015	
Shelly	5/25/2014	F	Joe	8/22/1978	M	bio father	5/25/2014	
Tom	11/19/2011	M	Joe	8/22/1978	M	bio father	11/19/2011	
Meg	7/19/2015	F	Sue	1/11/1983	F	bio mother	7/19/2015	
Shelly	5/25/2014	F	Sue	1/11/1983	F	bio mother	5/25/2014	
Tom	11/19/2011	M	Sue	1/11/1983	F	bio mother	11/19/2011	

- A new record is created for each primary client listing the other sibling as a secondary client to record the **sibling relationship type.**

Primary Client	Primary DOB	Primary Gender	Secondary Client	Secondary DOB	Secondary Gender	S:P Relation	Start	End
Shelly	5/25/2014	F	Meg	4/19/2015	F	M sibling	7/19/2015	
Shelly	5/25/2014	F	Tom	11/19/2011	M	M sibling	5/25/2014	
Shelly	5/25/2014	F	Tom	11/19/2011	M	P sibling	5/25/2014	
Tom	11/19/2011	M	Meg	4/19/2015	F	M sibling	7/19/2015	
Tom	11/19/2011	M	Shelly	5/25/2014	F	M sibling	5/25/2014	
Tom	11/19/2011	M	Shelly	5/25/2014	F	P sibling	5/25/2014	
Meg	7/19/2015	F	Shelly	5/25/2014	F	M sibling	7/19/2015	
Meg	7/19/2015	F	Tom	11/19/2011	M	P sibling	7/19/2015	

Step 4

Add Couple Relationships

- Marriage and cohabitation data are not available in the CHHS data files, so **identifying married couples and cohabitating unions** can be challenging.
- Individuals who share common children, however, could be identified as **“Parenting Couples.”**

Primary Client	Primary DOB	Primary Gender	Secondary Client	Secondary DOB	Secondary Gender	S:P Relation	Start	End
Shelly	5/25/2014	F	SUe	1/11/1983	F	bio mother	5/25/2014	
Shelly	5/25/2014	F	Joe	8/22/1978	M	bio father	5/25/2014	
Shelly	5/25/2014	F	Tom	11/19/2011	M	M sibling	5/25/2014	
Shelly	5/25/2014	F	Tom	11/19/2011	M	P sibling	5/25/2014	
Shelly	5/25/2014	F	Meg	4/19/2015	F	M sibling	7/19/2015	
Sue	1/11/1983	F	Shelly	5/25/2014	F	bio daughter	5/25/2014	
Sue	1/11/1983	F	Tom	11/19/2011	M	bio son	11/19/2011	
Sue	1/11/1983	F	Meg	7/19/2015	F	bio daughter	7/19/2015	
Sue	1/11/1983	F	Joe	8/22/1978	M	COUPLE	5/25/2014	
Sue	1/11/1983	F	Frank	11/19/1981	M	COUPLE	7/19/2015	
Joe	8/22/1978	M	Shelly	5/25/2014	F	bio daughter	5/25/2014	
Joe	8/22/1978	M	Tom	11/19/2011	M	bio son	11/19/2011	
Joe	8/22/1978	M	Sue	1/11/1983	F	COUPLE	5/25/2014	
Tom	11/19/2011	M	Sue	1/11/1983	F	bio mother	11/19/2011	
Tom	11/19/2011	M	Joe	8/22/1978	M	bio father	11/19/2011	
Tom	11/19/2011	M	Shelly	5/25/2014	F	M sibling	5/25/2014	
Tom	11/19/2011	M	Shelly	5/25/2014	F	P sibling	5/25/2014	
Tom	11/19/2011	M	Meg	4/19/2015	F	M sibling	7/19/2015	
Meg	7/19/2015	F	Sue	1/11/1983	F	bio mother	7/19/2015	
Meg	7/19/2015	F	Frank	11/19/1981	M	bio father	7/19/2015	
Meg	7/19/2015	F	Shelly	5/25/2014	F	M sibling	7/19/2015	
Meg	7/19/2015	F	Tom	11/19/2011	M	P sibling	7/19/2015	
Frank	11/19/1981	M	Meg	7/19/2015	F	bio daughter	7/19/2015	
Frank	11/19/1981	M	Sue	1/11/1983	F	COUPLE	7/19/2015	

Step 5

Identify Family Units

- Relationship file is merged with the person file to define all relationships for the the focal client.

Primary Client	Primary DOB	Primary Gender	Secondary Client	Secondary DOB	Secondary Gender	S:P Relation	Start	End
Shelly	5/25/2014	F	Shelly	5/25/2014	F	SELF	5/25/2014	
Shelly	5/25/2014	F	Sue	1/11/1983	F	bio mother	5/25/2014	
Shelly	5/25/2014	F	Joe	8/22/1978	M	bio father	5/25/2014	
Shelly	5/25/2014	F	Tom	11/19/2011	M	M sibling	5/25/2014	
Shelly	5/25/2014	F	Tom	11/19/2011	M	P sibling	5/25/2014	
Shelly	5/25/2014	F	Meg	4/19/2015	F	M sibling	7/19/2015	
Sue	1/11/1983	F	Sue	1/11/1983	F	SELF	1/11/1983	
Sue	1/11/1983	F	Shelly	5/25/2014	F	bio daughter	5/25/2014	
Sue	1/11/1983	F	Tom	11/19/2011	M	bio son	11/19/2011	
Sue	1/11/1983	F	Meg	7/19/2015	F	bio daughter	7/19/2015	
Joe	8/22/1978	M	Joe	8/22/1978	M	SELF	8/22/1978	
Joe	8/22/1978	M	Shelly	5/25/2014	F	bio daughter	5/25/2014	
Joe	8/22/1978	M	Tom	11/19/2011	M	bio son	11/19/2011	
Tom	11/19/2011	M	Tom	11/19/2011	M	SELF	11/19/2011	
Tom	11/19/2011	M	Sue	1/11/1983	F	bio mother	11/19/2011	
Tom	11/19/2011	M	Joe	8/22/1978	M	bio father	11/19/2011	
Tom	11/19/2011	M	Shelly	5/25/2014	F	M sibling	5/25/2014	
Tom	11/19/2011	M	Shelly	5/25/2014	F	P sibling	5/25/2014	
Tom	11/19/2011	M	Meg	4/19/2015	F	M sibling	7/19/2015	
Meg	7/19/2015	F	Meg	7/19/2015	F	SELF	7/19/2015	
Meg	7/19/2015	F	Sue	1/11/1983	F	bio mother	7/19/2015	
Meg	7/19/2015	F	Frank	11/19/2011	M	bio father	7/19/2015	
Meg	7/19/2015	F	Shelly	5/25/2014	F	M sibling	7/19/2015	
Meg	7/19/2015	F	Tom	11/19/2011	M	P sibling	7/19/2015	
Frank	11/19/1981	M	Meg	7/19/2015	F	bio daughter	7/19/2015	
Frank	11/19/1981	M	Sue	1/11/1983	F	COUPLE	7/19/2015	

Family Unit - Record Layout

- This roster can then be used **to identify family relationships** from the perspective of the focal client at a particular point in time.

Primary Client	Primary DOB	Primary Gender	Secondary Client	Secondary DOB	Secondary Gender	S:P Relation	Start	End
Shelly	5/25/2014	F	Shelly	5/25/2014	F	SELF	5/25/2014	
Shelly	5/25/2014	F	Sue	1/11/1983	F	bio mother	5/25/2014	
Shelly	5/25/2014	F	Joe	8/22/1978	M	bio father	5/25/2014	
Shelly	5/25/2014	F	Tom	11/19/2011	M	M & P sibling	5/25/2014	
Shelly	5/25/2014	F	Meg	4/19/2015	F	M sibling	7/19/2015	

Step 6

Assign Family Structure

- The information from the roster can be used to **assign general family structures** – such as single parent families, two-biological-parent families and families with complex sibling (half) compositions.

Parent Type	Family Type	SELF	COUPLE	Bio Mother	Bio Father	Whole siblings	Half M or P Siblings
Mother only	Mother only - 1 child no siblings	X		X			
	Mother only, whole siblings	X		X		X	
	Mother only, whole and half siblings	X		X		X	X
	Mother only, half siblings	X		X			X
Father only	Father only - 1 child no siblings	X			X		
	Father only, whole siblings	X			X	X	
	Father only, whole and half siblings	X			X	X	X
	Father only, half siblings	X			X		X
Two parents	Two parents -1 child no siblings	X	X	X	X		
	Two parents, whole siblings	X	X	X	X	X	
	Two parents, whole and half siblings	X	X	X	X	X	X
	Two parents, half siblings	X	X	X	X		X
Couple	No children	X	X				
Single	No children	X					

Advantages

- It is possible to create relatively **detailed biological and coupling relationship information** using administrative data.
- Information quality can be **improved** as additional data sources are added.

Limitations

- **Cannot determine** whether identified families actually **reside together**.
- Doing so would require a **blended approach** whereby the family units are cross-referenced with the household file records, to examine which family members reside in a household at a particular point and time.
- **Further exploration** of this method is suggested once the strategies for constructing basic household and family level units are successfully implemented.

Conclusion

- With exceptionally **broad coverage** of CHHS service populations (and multiple sources of address and relationship data), the integrated data from the CHHS Record Reconciliation Project provides a unique opportunity to link individuals at the household- and family-level improving policy and program research.
- Household and Family linkage keys and structural classifications could be **added to the CHHS file for each individual record**.
- The key could be utilized by agencies conducting **cross program research**.
- Keys could also be used to **improve individual level linkages** by incorporating family information as a matching field.

Questions?

For further information on the CHHS Record Reconciliation Project

<https://www.datanetwork.org/research/chhs-record-reconciliation-and-research-data-hub/>

Contact

Regan Foust

rfoust@usc.edu

Stephanie Cuccaro-Alamin

stephca@berkeley.edu