

# Using Synchronized Contact Information in Multi-Mode CAPI Projects

Kate Krueger

University of Wisconsin Survey Center  
University of Wisconsin-Madison

International Field Directors and Technologies Conference  
New Orleans, LA  
May 22, 2017

© 2016. Materials may not be reproduced without permission of the author.



# Overview

---

- Project Operations Database (POD) and CAPI
- Mailsync
- Data duplication
- Grouping households
- Looking to the future



# Overview

---

- Project Operations Database (POD) and CAPI
- Mailsync
- Data duplication
- Grouping households
- Looking to the future



# Project Operations Database (POD)

---

- Mail, Web, and CAPI (In-person) Studies
- Manage data about sample members' contact information
- Record and track mailing attempts
- Access Database Front-end
- MySQL tables - Hierarchical model



# Contact Management in CAPI Studies

---

- Headquarters
  - Mailings: Advance Letter, Thank You Postcard, Follow-up SAQ
  - CATI lab: Follow-up phone interview
- In the field (CAPI Interviewers)
  - Telephone calls
  - In-person visits



# Contact Information

---

- Telephone
- Address
- Email
- Person



# The Old Days

---

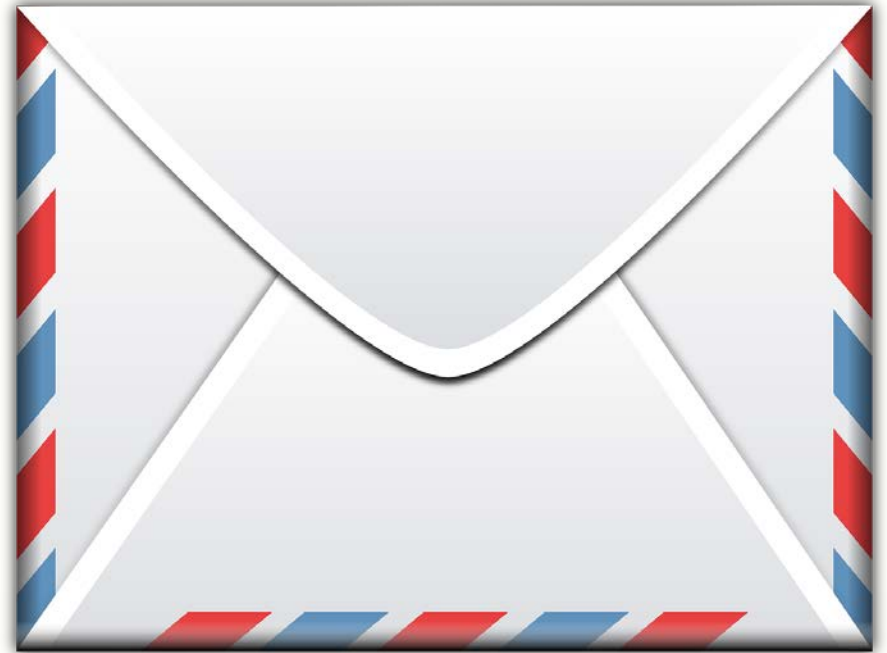
- Flat format in CASES
  - Only 1 “Home Address”, only 1 “Work Address”, etc
- Tracer
  - 4 separate tables: telephone, address, email, person
  - Collected historical information
  - Still restricted by flat format of CASES



# Data Modification in the Field and at Headquarters

---

- CaseID 10000
- Advance Letter sent out







# Data Modification in the Field and at Headquarters

---

- CaseID 10000
- Advance Letter sent out
- Interviewer begins working the case: telephone calls, knocking on doors





# Data Modification in the Field and at Headquarters

---

- CaseID 10000
- Advance Letter sent out
- Interviewer working the case
- Advance letter sent back undeliverable
- Tracking and Locating Department traces case and adds new address





# Data Modification in the Field and at Headquarters

---

- CaseID 10000
- Advance Letter sent out
- Interviewer working the case
- Advance letter sent back undeliverable
- Tracking and Locating Department traces case and adds new address
- **Contact the old address?**





# Transitioning to Hierarchical Model in CAPI

---

- Rosters within a case
  - Address
  - Telephone
  - Email
  - Person
- Multiple people and their contact information in a case



# Overview

---

- Project Operations Database (POD) and CAPI
- **Mailsync**
- Data duplication
- Grouping households
- Looking to the future



# Mailsync

---

- Asynchronous data transfer program
- Executed on the CAPI laptop during synchronization process
- Import and export into POD (headquarters) and CASES (field)
  - MySQL tables  $\leftrightarrow$  CASES rosters
- Data Collision
  - CASES evaluates based on timestamps in CASES and POD



# Mailsync: Importing Data from POD to CASES

---

1. Executes a Perl script that opens each of the 4 MySQL tables and writes out relevant data to text files
2. Runs CASES using Batch Mode and imports the data from the text file using CSMGET
3. The data imported is appended/updated in 4 respective CASES rosters



# Mailsync: Exporting Data from CASES to POD

---

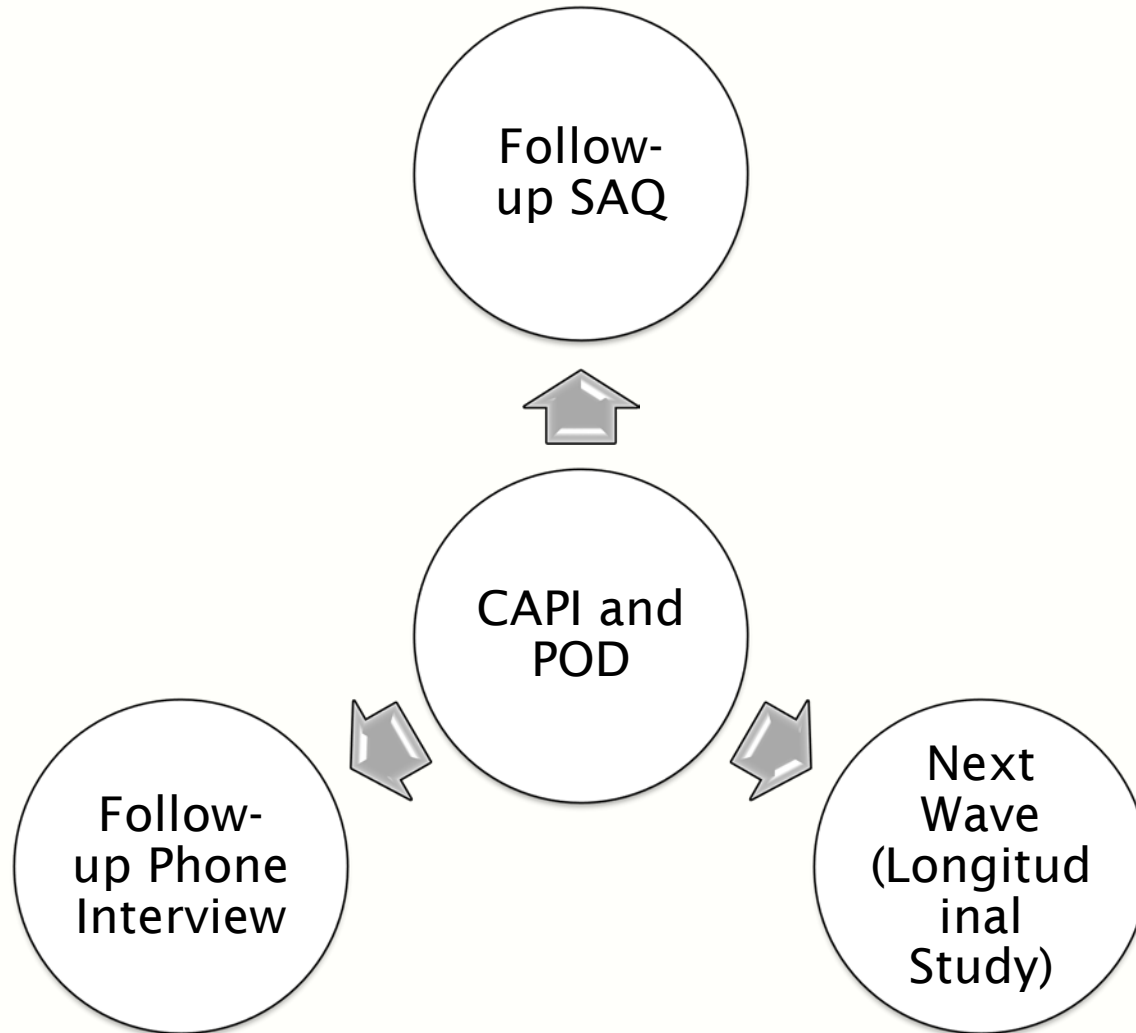
1. CASES writes out relevant data from 4 CASES rosters to text files
2. Executes a Perl script to read text files into 4 MySQL tables





# Multi-Mode Studies

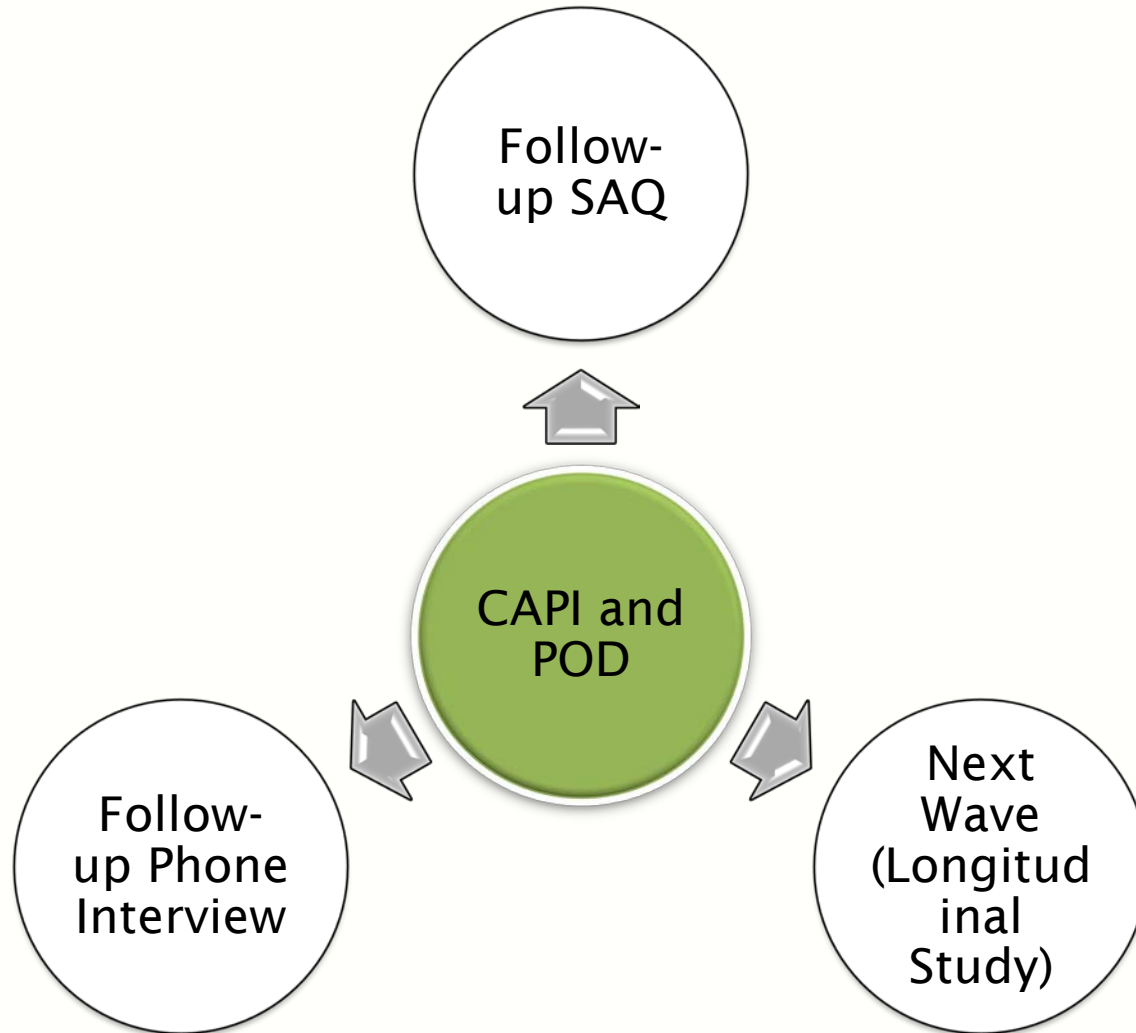
---





# Multi-Mode Studies

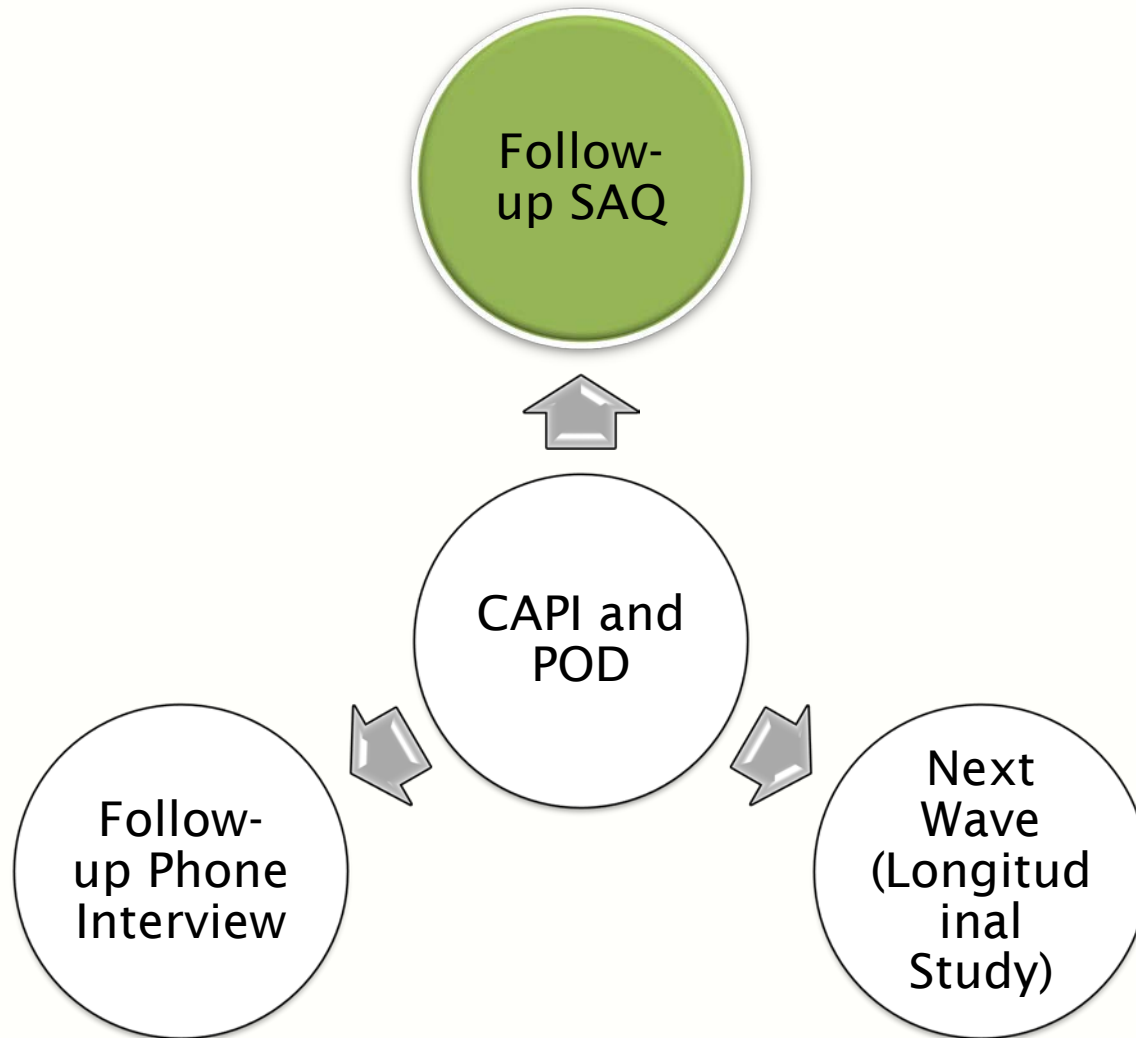
---





# Multi-Mode Studies

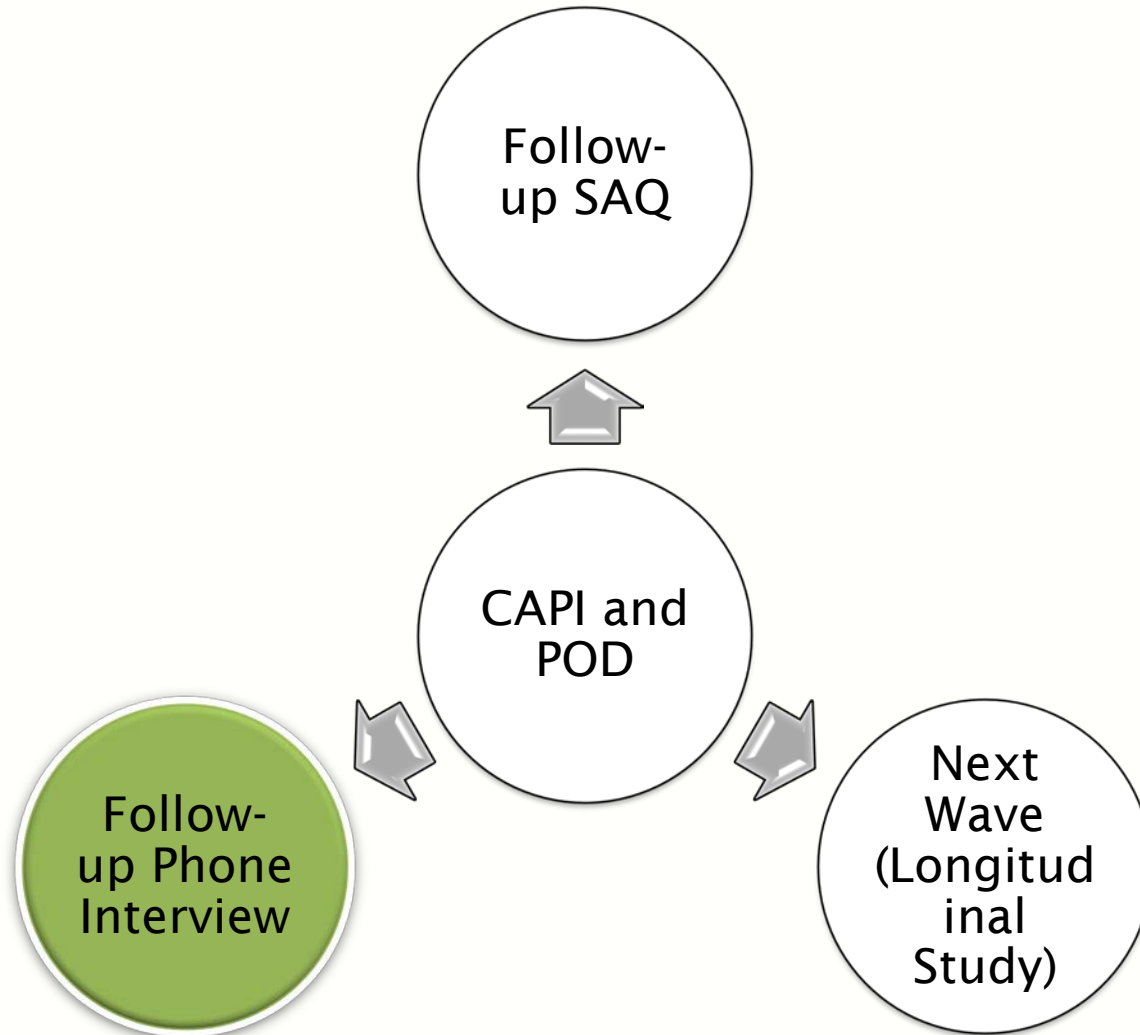
---





# Multi-Mode Studies

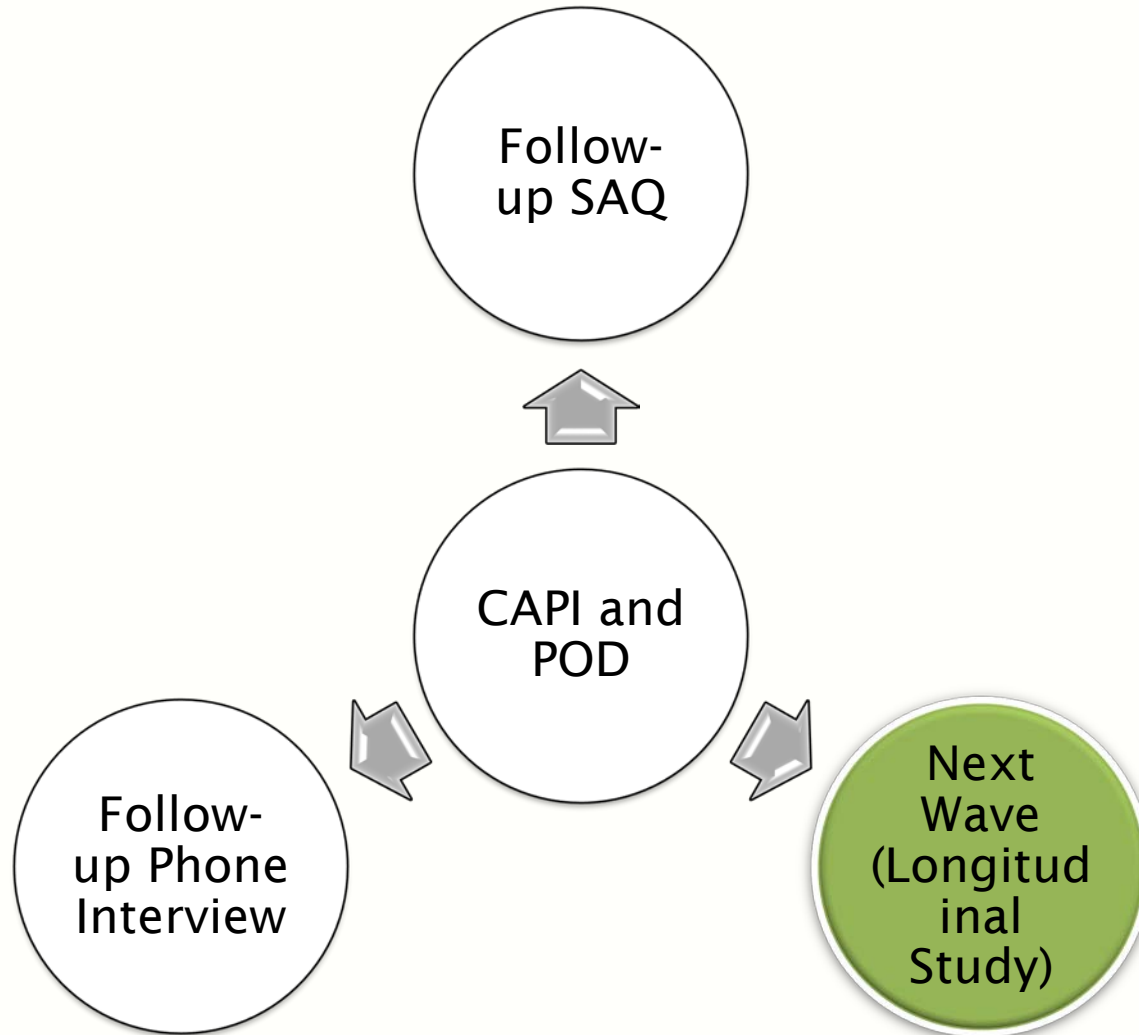
---





# Multi-Mode Studies

---





# Overview

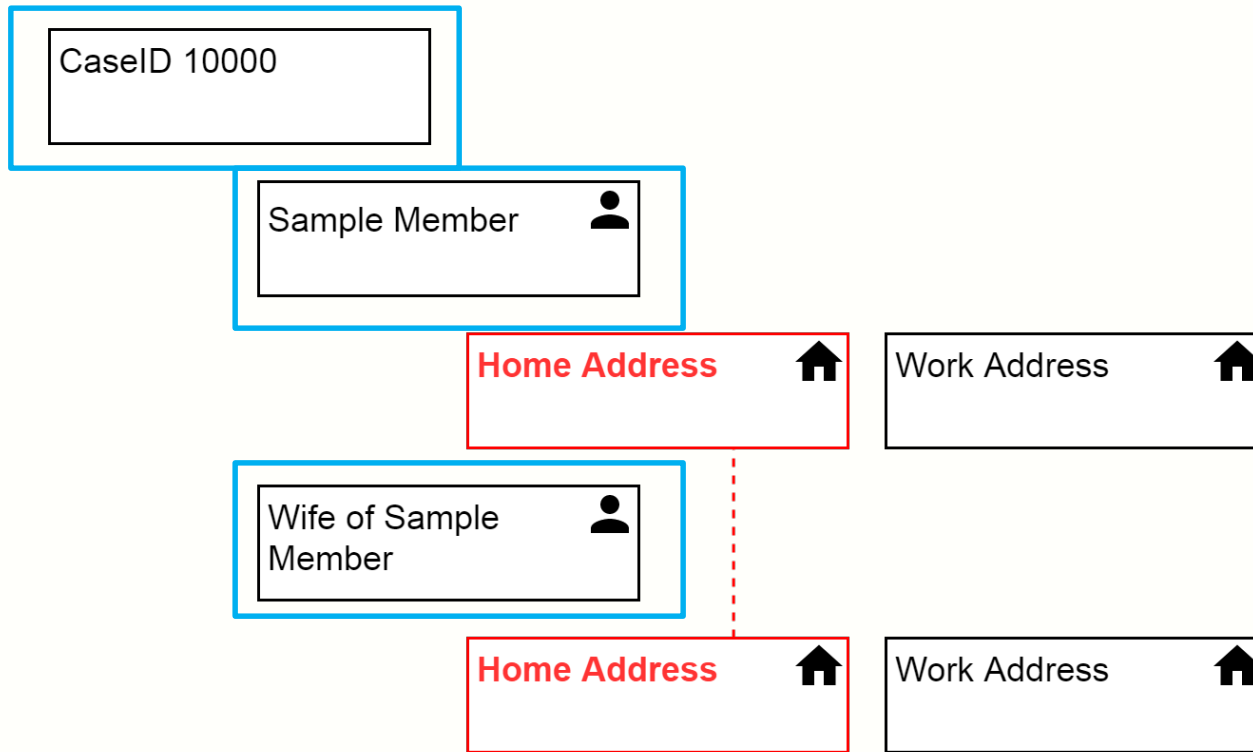
---

- Project Operations Database (POD) and CAPI
- Mailsync
- **Data duplication**
- Grouping households
- Looking to the future



# Data Duplication Within Case (Multiple People)

---





# Data Duplication within a case

---

- Why is it a problem?
  - Increased complexity for interviewers and tracers
  - More coding and cleaning
  - Duplicate data = duplicated effort
- Solutions to data duplication?
  - Identify them?
  - Avoid them?



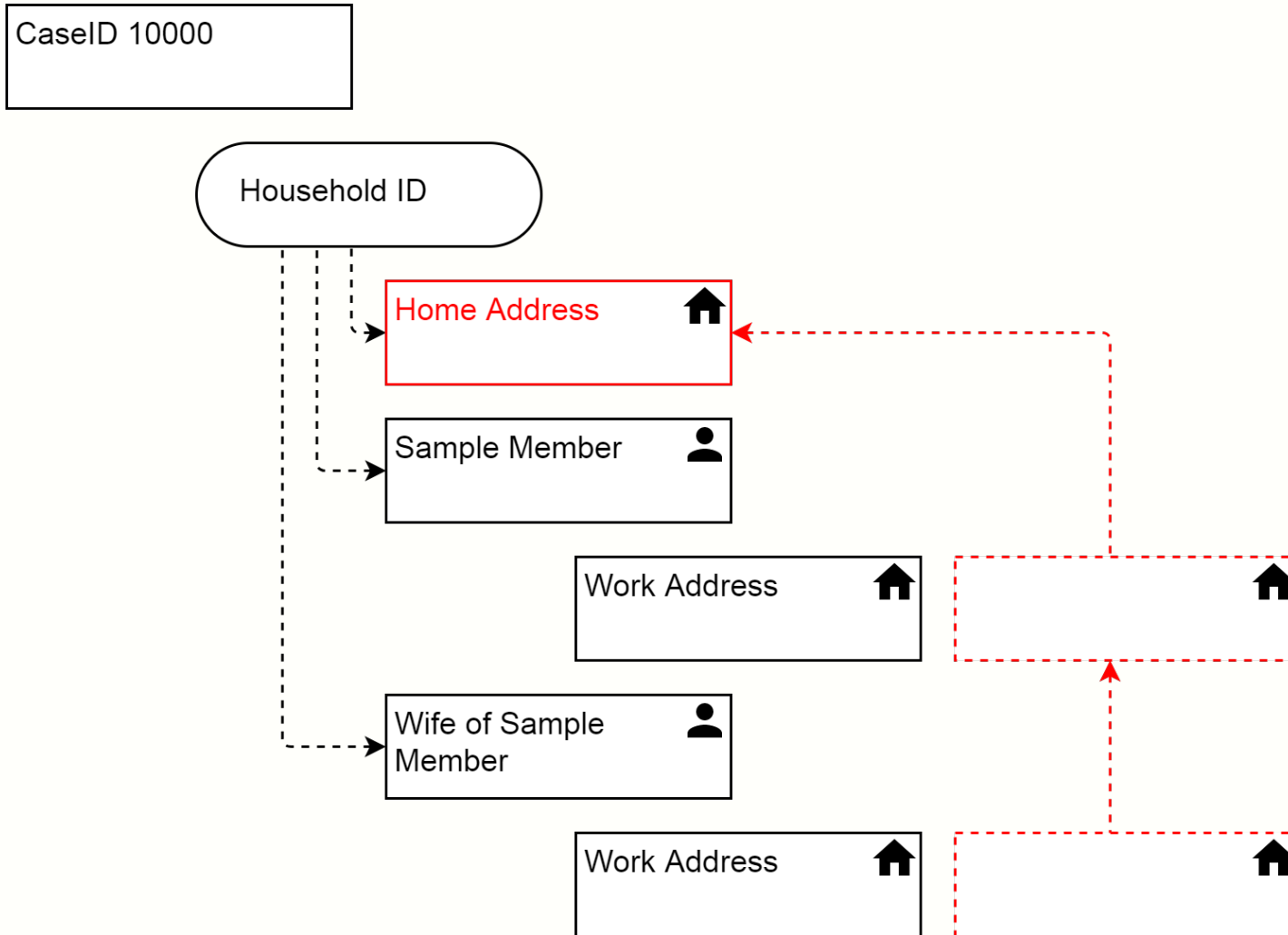


# Overview

---

- Project Operations Database (POD) and CAPI
- Mailsync
- Data duplication
- **Grouping households**
- Looking to the future

# 1<sup>st</sup> Implementation: Grouping by Household ID





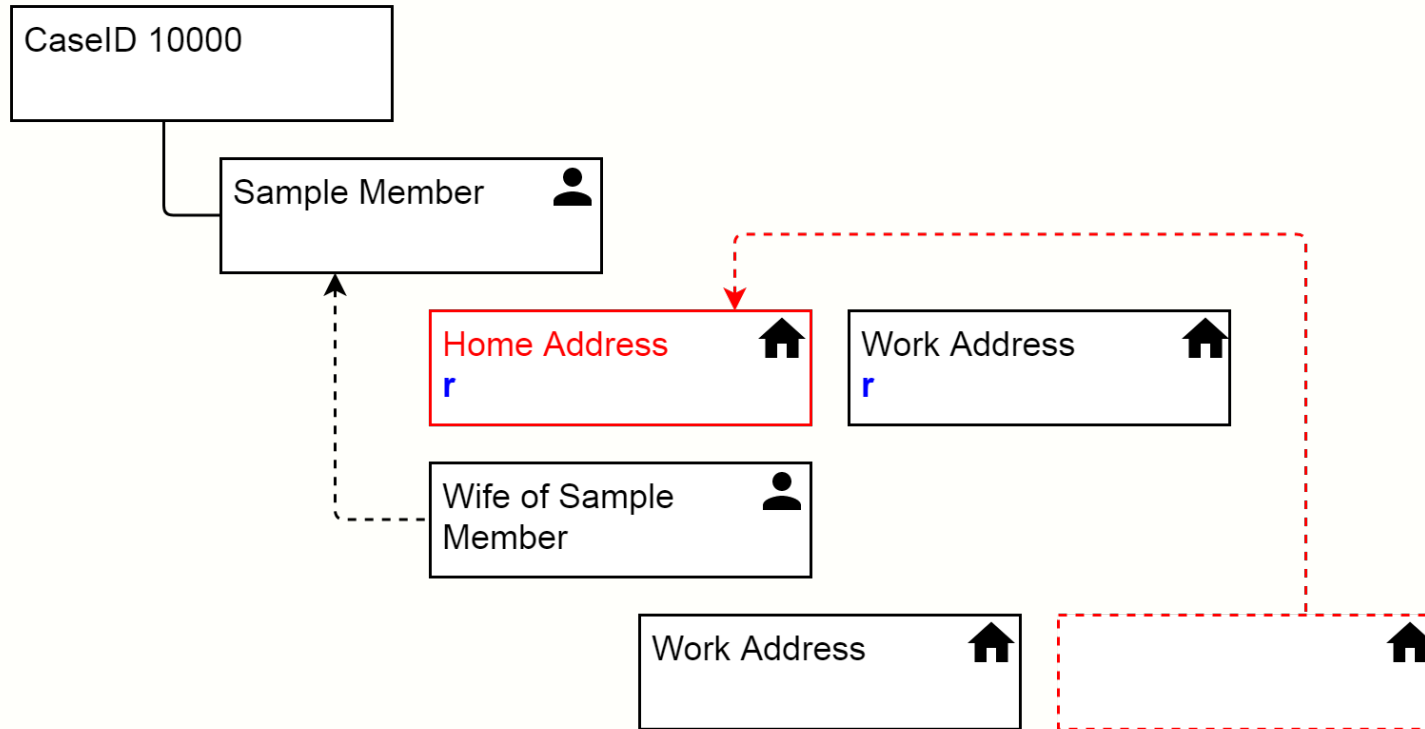
# Grouping by Household ID

---

- Relies on household enumeration
  - No household or kinship identifiers in sample
  - POD users cannot enumerate household
- Increased complexity for tracers
  - Arbitrary “household ids”
- Incompatible with hierarchical data structure
  - Priority given to household rather than sample member



# Current Implementation: Grouping by Individual





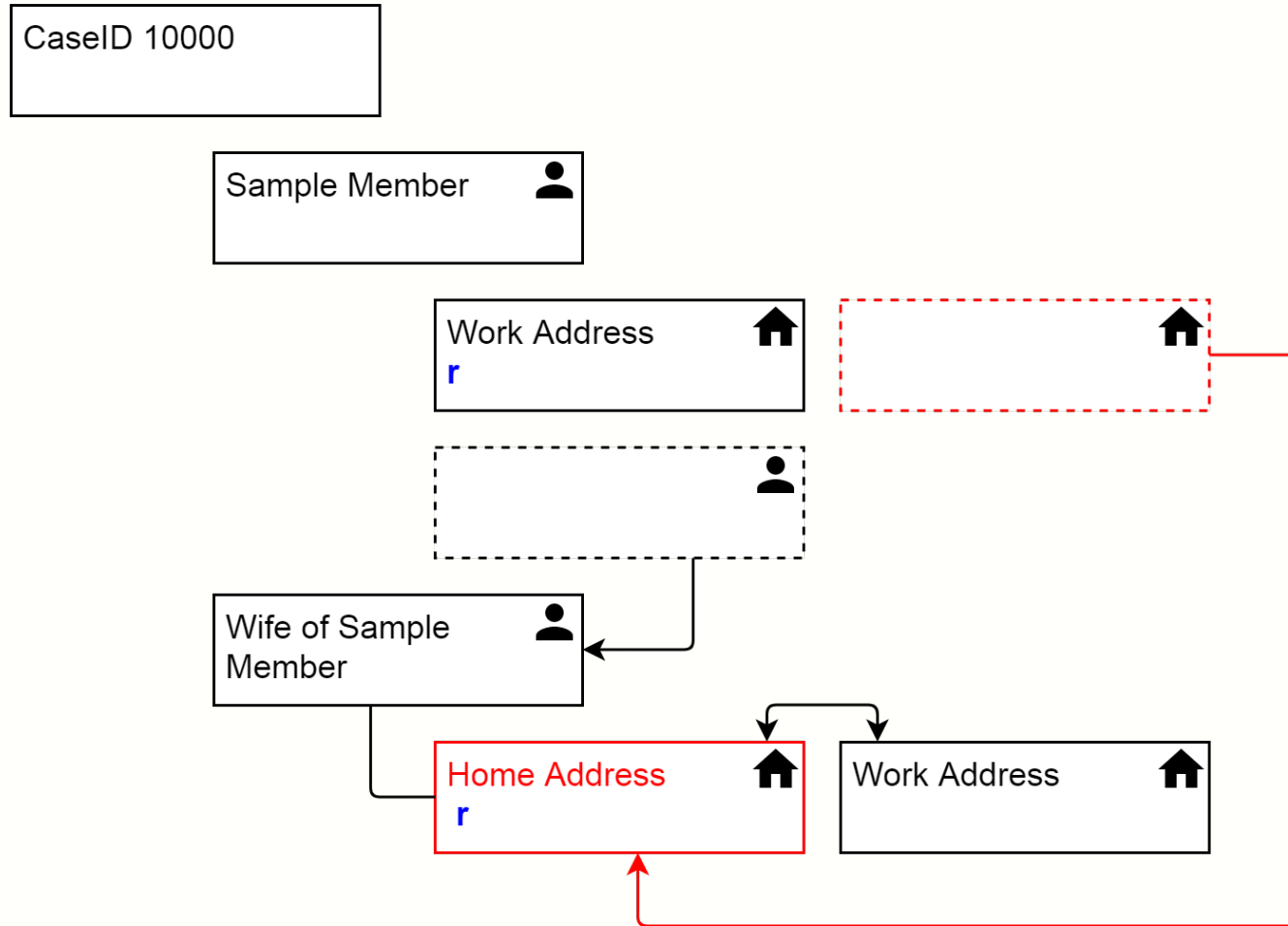
# Grouping by Individual

---

- Does not rely on household enumeration
- Decreased complexity for tracers
  - No need to manage “household identifiers”
- Compatible with hierarchical person structure
  - Priority given to sample member



# Grouping by Individual





# Disadvantages of Grouping

---

- Grouping is a reflection of one point in time
- Increased complexity for interviewers
  - Incorrect groupings make a case harder to manage
- Loss of data about past residences
  - What if two people no longer share the address?
  - “Who gets to keep the house?”
  - Respondent (Sample Member) association flags



# Overview

---

- Project Operations Database (POD) and CAPI
- Mailsync
- Data duplication
- Grouping households
- **Looking to the future**





# Future Outlook

---

- Better displays for interviewers/POD users
- Is grouping useful?
- Non-hierarchical data structure?

Thank You!

For copies of this presentation or more information, contact:

Kate Krueger  
kgolen@ssc.wisc.edu

Please visit us at:  
**[www.uwsc.wisc.edu](http://www.uwsc.wisc.edu)**