

# Examining Data Quality with Changing Patterns of Telephone Usage in California Health Interview Survey

**SHADAC Workshop:  
Issues in Survey Methods for State Health Surveys**  
January 13, 2009

Sunghee Lee  
UCLA Center for Health Policy Research



## *Today's presentation*

- ▶ **California Health Interview Survey (CHIS)**
- ▶ **CHIS 2007 methodological studies**
  - ▶ Background
  - ▶ Nonresponse study
  - ▶ Noncoverage study
    - ▶ Cell-only population
    - ▶ Cell-mainly population
    - ▶ Weighting survey data with cell-phone samples



## ***California Health Interview Survey***



### ***California Health Interview Survey (CHIS)***

- ▶ **UCLA Center for Health Policy Research**
- ▶ **Conducted every 2 years since 2001**
- ▶ **Covers a broad range of health-related topics**
- ▶ **Target population: All Californians living in households, excluding persons living in “group quarters” and homeless**
- ▶ **Conducted in five languages: English, Spanish, Chinese (Mandarin, Cantonese), Vietnamese, Korean**



4

## *CHIS: Sample design – 1*

- ▶ **Geographically stratified random digit dial (RDD) telephone survey designed to provide statistically reliable estimates:**
  - ▶ At the local level for counties and statewide
  - ▶ For major racial/ethnic groups and many ethnic subgroups
- ▶ **Sample design**
  - ▶ Stratified simple random sampling
  - ▶ 58 counties are designed to represent 44 total strata
    - ▶ 41 individual county strata
    - ▶ 17 smallest counties combined into 3 grouped county strata
  - ▶ Substratification for oversampling ethnic groups



5

## *CHIS: Sample design – 2*

- ▶ **Random digit dialing computer-assisted telephone interviews**
- ▶ **One adult selected at random in each household, plus children and teens if available**
- ▶ **Three separate interviews**
  - ▶ Adult (age 18+) ~ 30 minutes
  - ▶ Adolescents (age 12-17) ~ 20 minutes
  - ▶ Children (age 0-11, by adult proxy) ~ 15 minutes



6

## *CHIS: Sample size*

	2001	2003	2005	2007
Adult (18+)	56,270	42,044	43,020	51,048
Child (0-11)	12,802	8,526	11,358	9,918
Teen (12-17)	5,733	4,010	4,029	3,638



7

## ***CHIS 2007 Methodological Studies***



## ***Background***



## ***Background***

- ▶ **Changes to the survey environment**
  - ▶ Declining response rates
    - Increase in the potential for nonresponse bias
- ▶ **Changes to the communication channels**
  - ▶ Growth of cell-phone only households
    - Increase in the potential for noncoverage bias in landline RDD sample frames
  - ▶ Growth of answering machine and caller ID
    - Decrease in response rates; increase in nonresponse bias
- **Threatens data representativeness and quality**
  
- ▶ **CHIS is not an exception**



## ***Nonresponse Bias***



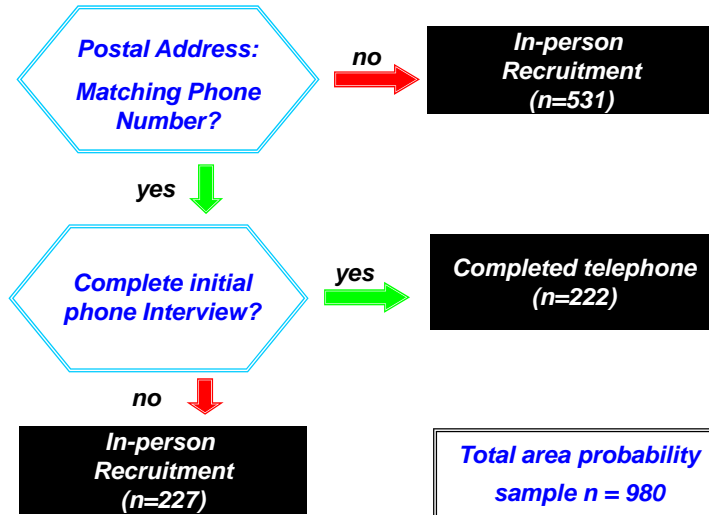
### ***Exploring nonresponse bias in CHIS 2007***

- ▶ **Telephone plus in-person**
  - ▶ Separate area probability sample to complement the RDD sample in Los Angeles County
    - ▶ Cost concerns of geographic dispersion of RDD nonresponders within a county
  - ▶ 12+ month planning process
  - ▶ Multiple “Methods Experts” planning teleconference calls
  - ▶ Individual consultation with experts in survey methodology
  - ▶ CHIS Advisory Board and Sample Design & Survey Methodology TAC



12

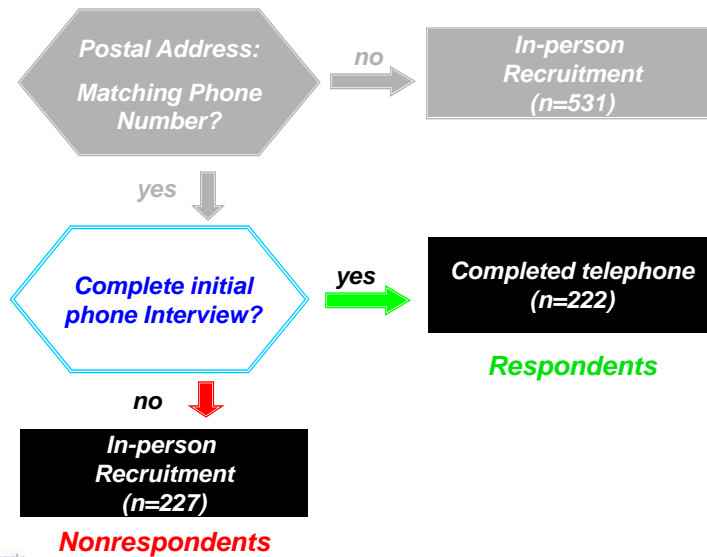
## CHIS 2007 area probability sample design



## CHIS 2007 nonresponse study

- ▶ Use data from the “matched” (telephone number to address) area probability sample
  - ▶ Comparisons over 50 variables
    - ▶ Telephone interview respondents (n=227)  
vs.
    - ▶ Combined telephone interview respondents and nonrespondents (n=449)
  - ▶ Analysis uses final weights adjusting at the total area probability sample level

## CHIS 2007 area probability sample design



## CHIS 2007 nonresponse study summary

- ▶ Few estimates with potential nonresponse biases
- ▶ Using respondents may overestimate
  - ▶ Non-Latino White
  - ▶ US-born citizen
  - ▶ Higher income
  - ▶ Having usual source of care
  - ▶ Visiting doctors past year
  - ▶ Health insurance coverage
- ▶ Using respondents may overestimate
  - ▶ Having STD test done past year

## *Noncoverage Bias*

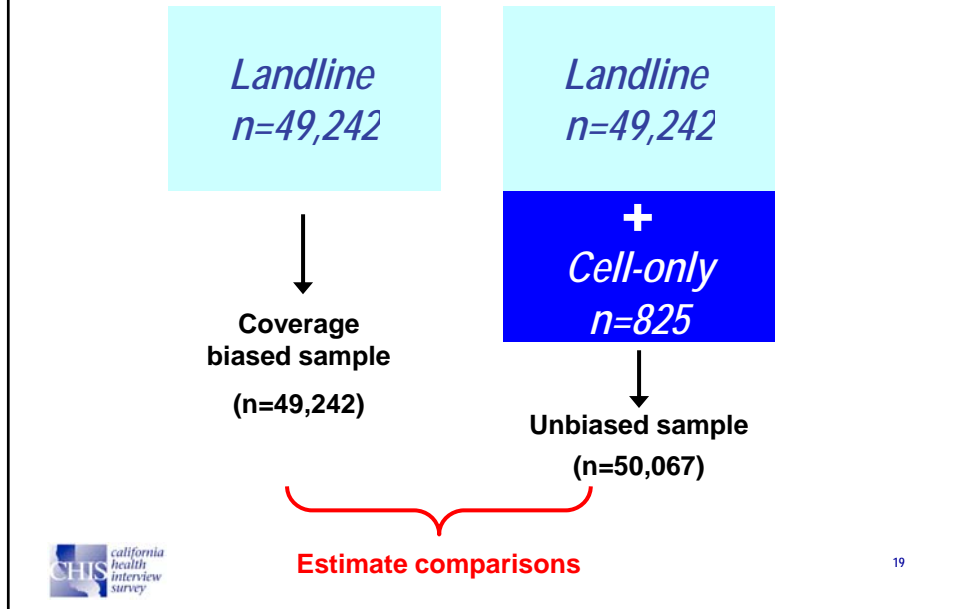


### *How to explore noncoverage in CHIS 2007?*

- ▶ **Evidence from NHIS that health/health related estimates differ by telephone status (Blumberg et al.)**
- ▶ **CHIS 2005: Pilot test of “cell-phone only” sample**
  - ▶ 98 completed
  - ▶ Successful in screening, reimbursing respondents for air time, and completing a full-length interview by cell phone  
(see Brick, Edwards, and Lee, *POQ* Special Issue 2007)
- ▶ **CHIS 2007: Expanded the cell-only sample for a target sample size of 800**

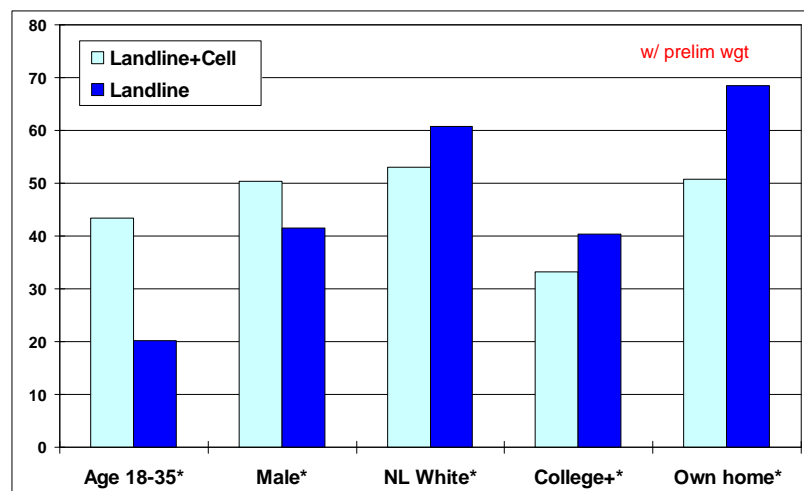


## How to explore noncoverage in CHIS 2007?



## CHIS 2007 noncoverage study findings

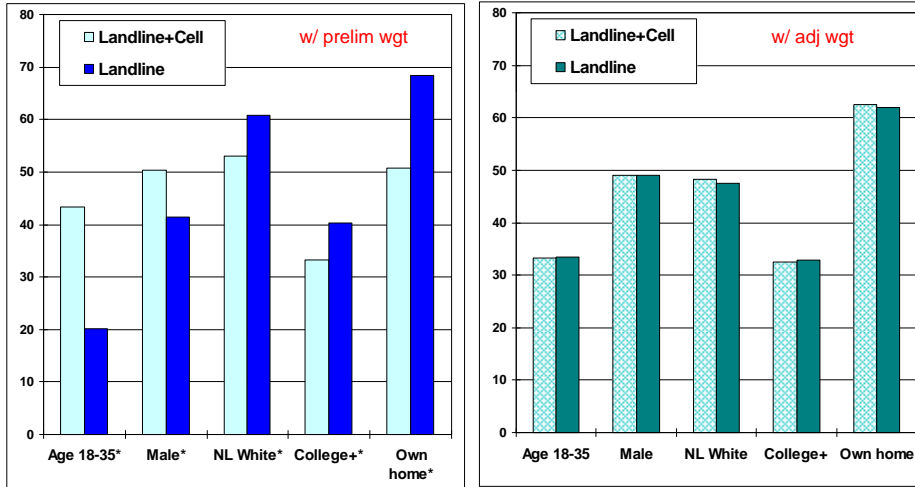
### ► Demographics controlled in adjustments



\* Statistically sig. 20

## CHIS 2007 noncoverage study findings

### Demographics in adjustments



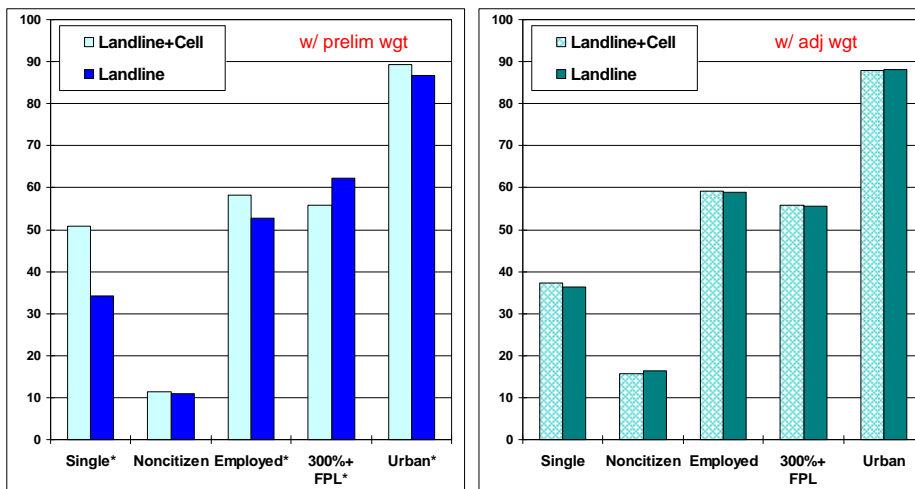
\* Statistically sig.



21

## CHIS 2007 noncoverage study findings

### Demographics not in adjustments



\* Statistically sig.



22

## *CHIS 2007 noncoverage study summary*

- ▶ **Landline RDD sample is more likely to overestimate**
  - ▶ Older
  - ▶ Female
  - ▶ Married
  - ▶ Home owner
  - ▶ With higher education and income
  
- ▶ **When adjustments are applied through weighting, the demographic and health condition, behavior and access differences disappear!!**



23

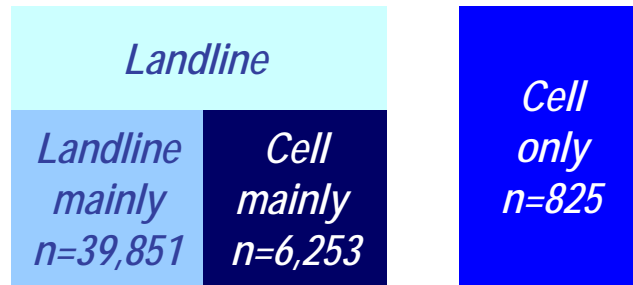
## *Weighting landline and cell-only data*

- ▶ **Multiple frame approach**
  - ▶ No overlap between landline RDD and cell-only samples
  - ▶ Prepare landline and cell-only base weight separately
    - ▶ Landline sample: the same procedure as previous
    - ▶ Cell-only sample: start with cell sample base weights and make cell-only adjustment assuming cell-only rate of 18% (NHIS 2007 West Region cell-only rate)
  - ▶ Combine landline and cell-only sample before raking
  - ▶ Apply raking
    - ▶ County, age, sex, race, ethnicity, education, # adults
    - ▶ Non-landline adjustment: age, education, home ownership, # adults



24

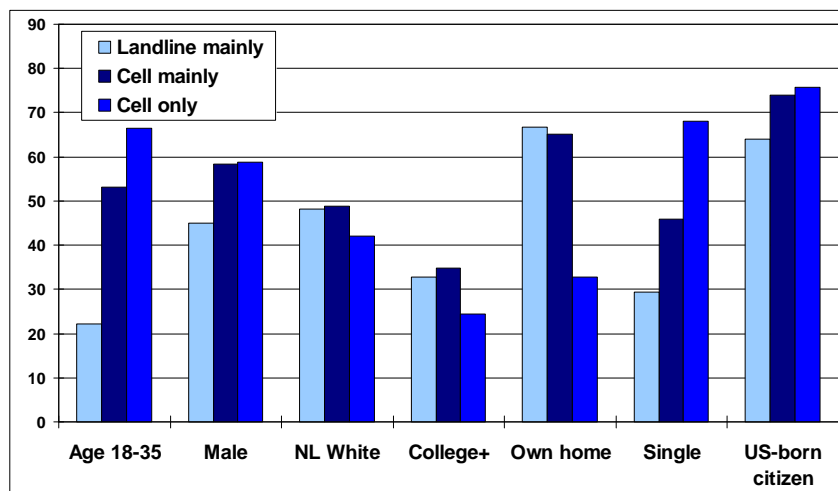
## How to explore cell-mainly population?



Comparisons by  
telephone usage

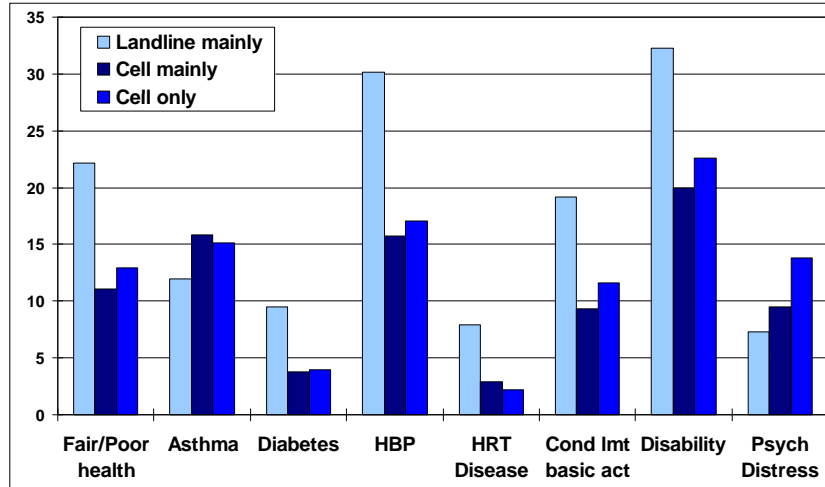
## CHIS 2007 noncoverage study findings

### Demographics



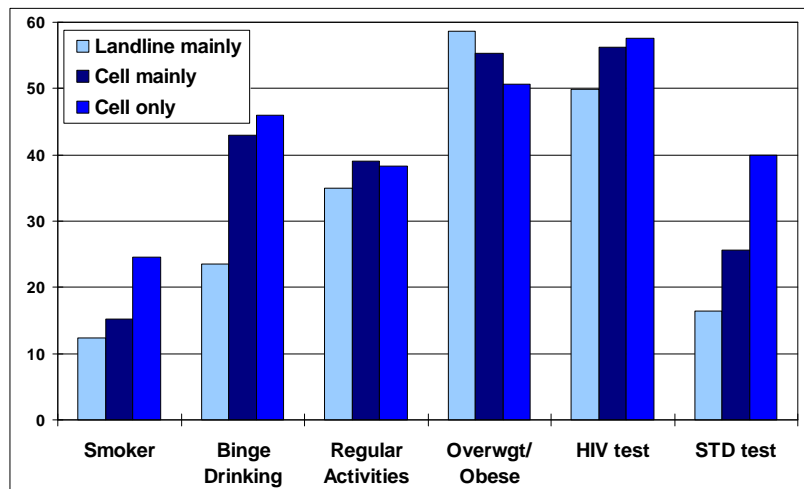
## CHIS 2007 noncoverage study findings

### ► Health conditions



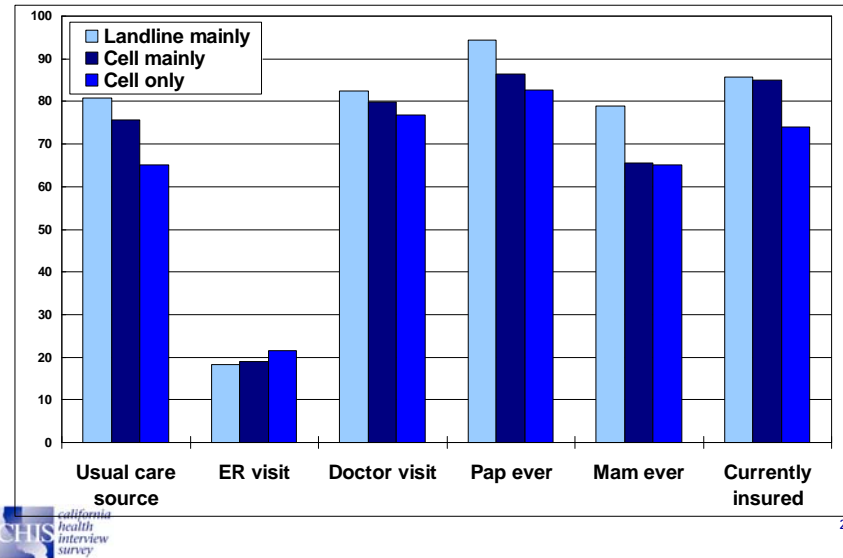
## CHIS 2007 noncoverage study findings

### ► Health risk behaviors



## CHIS 2007 noncoverage study findings

### ► Health care utilization and health insurance



## CHIS 2007 noncoverage study summary

- Traditional landline RDD sample is more likely to be
  - ▶ Older
  - ▶ Female
  - ▶ Married
  - ▶ Home owner
  - ▶ With higher education and income
- When adjustments are applied through weighting, the demographic and health condition, behavior and access differences disappear!!
- Cell-mainly are like neither cell-only nor landline-mainly

## *CHIS 2007 methodological study summary*

### ▶ **Nonresponse**

- ▶ Very few variables with bias (e.g., health insurance)
- ▶ However, the weight used in the study are not the best  
→ May expect further reduction in bias

### ▶ **Noncoverage**

- ▶ Weighting can effectively reduce bias
- ▶ However, cell-mainly may raise concerns
- ▶ Both cell-only and cell-mainly groups are rapidly growing

### ▶ **Cost implications**

- ▶ Noncoverage less expensive than nonresponse
- ▶ Potential cost saving by capturing cell-mainly in cell sample

