

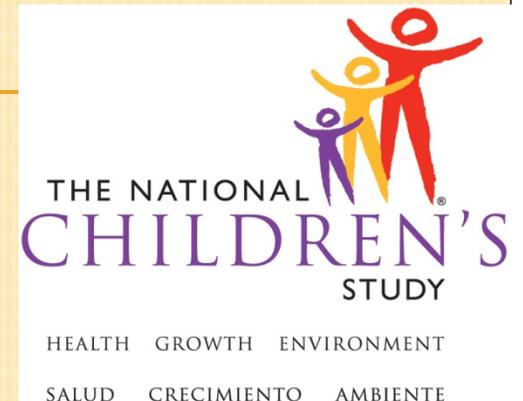
IMPLEMENTING PROVIDER-BASED SAMPLING FOR THE NATIONAL CHILDREN'S STUDY: OPPORTUNITIES AND CHALLENGES

SUMMARY OF A CONCEPT PAPER PREPARED BY AND AVAILABLE FROM THE NCS PROVIDER- BASED STUDY SITES

Prepared by: K. Belanger, S. Buka, D. Cherry, D. Dudley, D. Hale, I. Hertz-
Picciotto, N. Paneth, J. Robbins, E. Triche

On behalf of the NCS Provider-Based Study Sites

Stephen Buka, Sc.D.
Brown University
Providence RI Study Center



CENTRAL NCS OBJECTIVES THAT SHAPED ORIGINAL DESIGN

1. Study population should constitute a national probability sample
 2. Initiate study as early in child's life as possible
- × **These priorities led to initiation of household probability sampling as primary enrollment method**

OTHER NCS OBJECTIVES

1. Obtaining biological collections in first trimester
2. Making clinical assessments later in pregnancy
3. Obtaining extensive collections at delivery and after birth

These objectives might possibly have been achieved by enrolling from clinical sources.

PROBLEMS IMPLEMENTING HOUSEHOLD SAMPLING DESIGN

1. Difficult to encounter and enroll women in household during brief period of time—at most two months—when women aware they are in first trimester of pregnancy

Needle (pregnant women): Haystack (15,000 households)

2. Problems making arrangements with the birth hospitals for elaborate birth protocol, especially since some deliver only a handful of study babies

Resulted in recruitment challenges for first trimester pregnancies and incomplete collection of birth specimens

RESULTING NCS EXPERIMENT

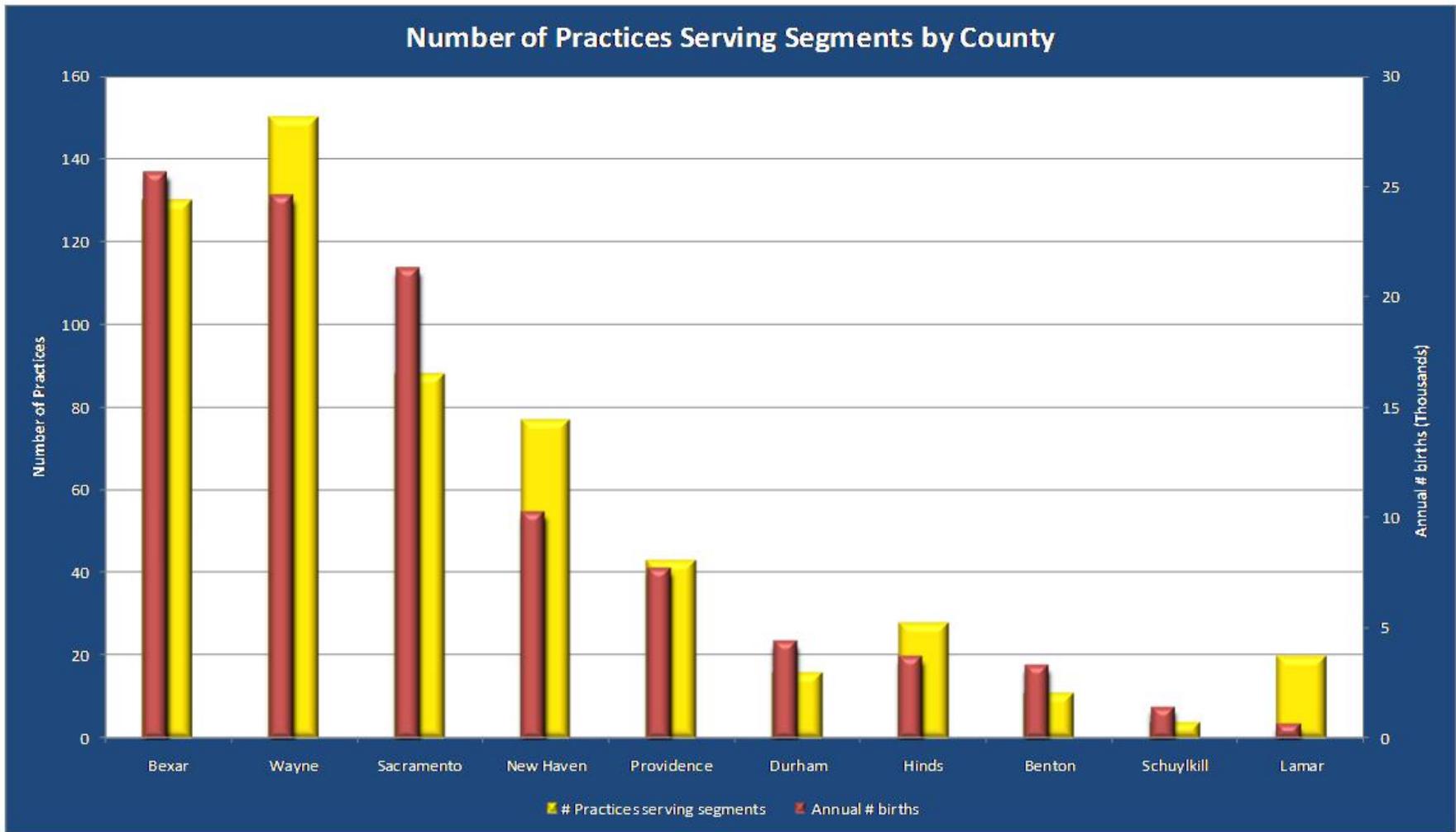
1. Three alternative forms of recruitment tested in 30 study centers (10 each), including provider-based recruitment
2. Household-based sampling retained in all three alternative recruitment strategies

Needle (eligible household): Haystack (pregnant women)

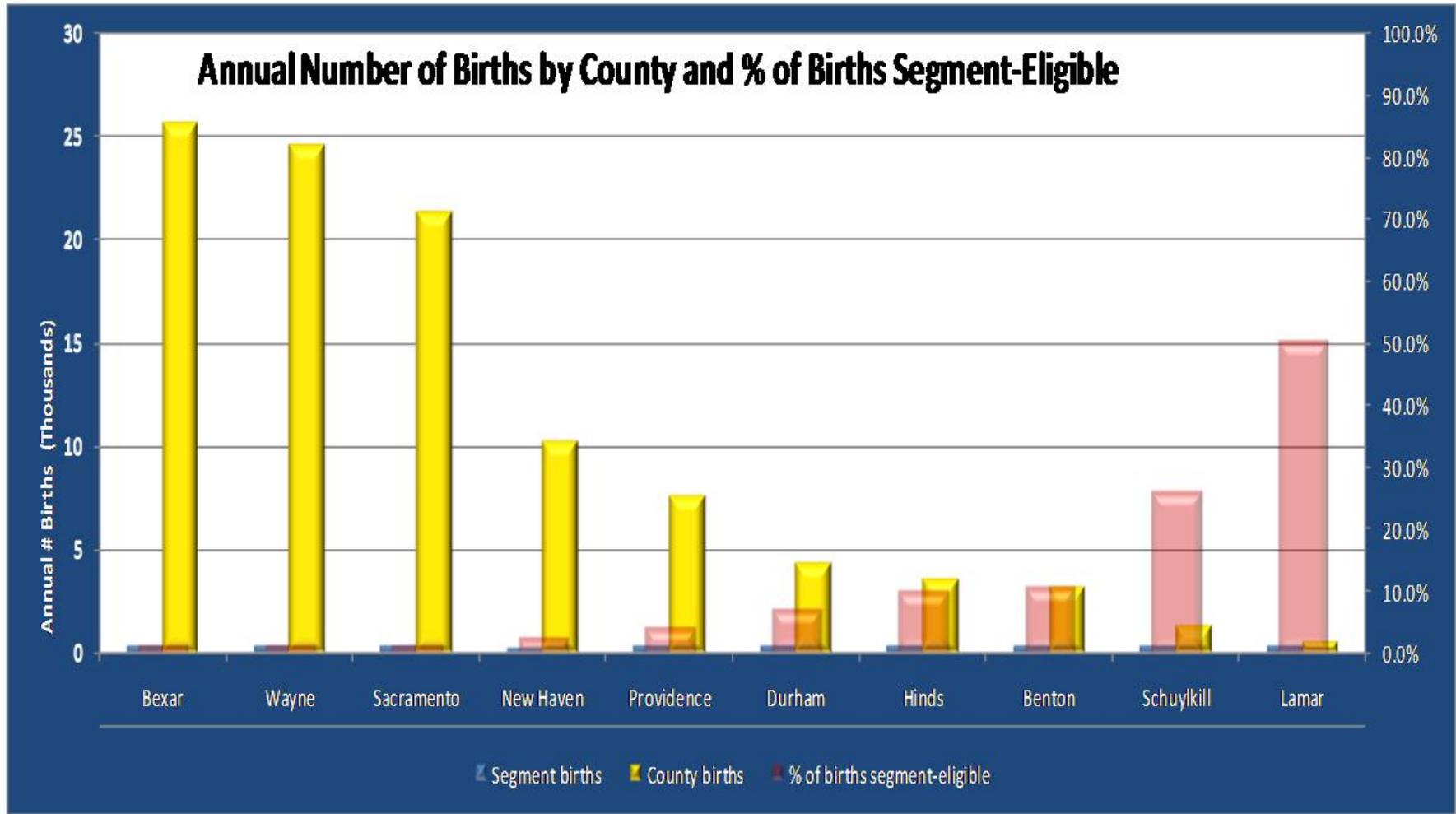
Women in prenatal care are address-matched to define eligibility; logistics; HIPAA; etc.

CENTRAL CHALLENGES WITH PROVIDER-BASED APPROACH

1. Considerable variability by study center in numbers of practices and proportion of births that are segment-eligible
2. In large counties, up to 150 prenatal care practices and up to 20-30 hospitals provide care to geographically dispersed, segment-eligible women. Only 1-2% of provider's patients may be eligible.
3. Small counties, conversely, are able to work with a small number of providers to recruit large proportion of segment-eligible women
4. But small counties have considerable proportion of women receiving prenatal care outside of county.



**NUMBER OF PRACTICES SERVING SEGMENTS AND
NUMBER OF BIRTHS BY COUNTY**



NUMBER OF BIRTHS AND % OF SEGMENT ELIGIBLE BIRTHS BY COUNTY

PROPOSAL FOR PROVIDER-BASED SAMPLING

OVERVIEW OF BASIC APPROACH

1. Selecting a Sample of Provider Groups
2. Sampling Women within Selected Provider Groups
3. Enrolling women through community engagement
4. Enrolling women with late or no prenatal care
5. Enrolling preconception cohort
6. Determining sample size based on realistic response rates

1. SELECTING A SAMPLE OF PROVIDER GROUPS

- ✘ Assemble sampling frame of *provider groups* (*practices*)
 - + Birth certificate data for comprehensive listing of providers
 - + Group providers by practice
- ✘ Stratify provider groups to insure specific subgroups will be represented
 - + Small # of stratification variables may vary by county (based on homogeneity)
 - + Additionally stratify by geographic location?
- ✘ Sample provider groups using stratified sampling
 - + 10-20 groups may be sufficient even in large counties (Dr. Michael Elliot, U Michigan)
 - + Sampling proportional to size (anticipated # prenatal care patients) within strata to draw probability-based sample
 - + Replace refusals with randomly selected group within that stratum

2. SAMPLING WOMEN WITHIN SELECTED PROVIDER GROUPS

- ✘ Except in very small PSUs, # of pregnant women per selected provider group will outnumber NCS target for PSU; therefore sampling of women required
- ✘ Options:
 - + Probability-based: systematic sampling (every n th patient); selecting on randomly sampled days of the week or weeks of the year; geographic subsampling
 - + Representative sample of women within provider groups (not probability-based): screen all interested women and select women by computer algorithm to obtain representative sample

3. ENROLLING WOMEN THROUGH COMMUNITY ENGAGEMENT

- ✘ Using strict probability-based provider sampling, enroll women only if receive care through sampled providers
- ✘ For representative non-probability based approach, community engagement could supplement provider group recruitment

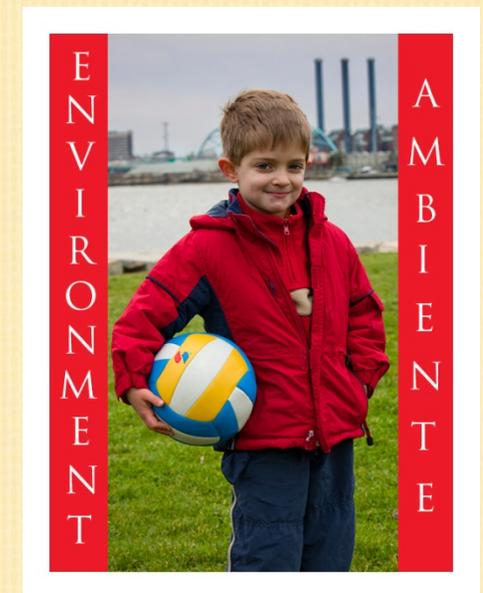
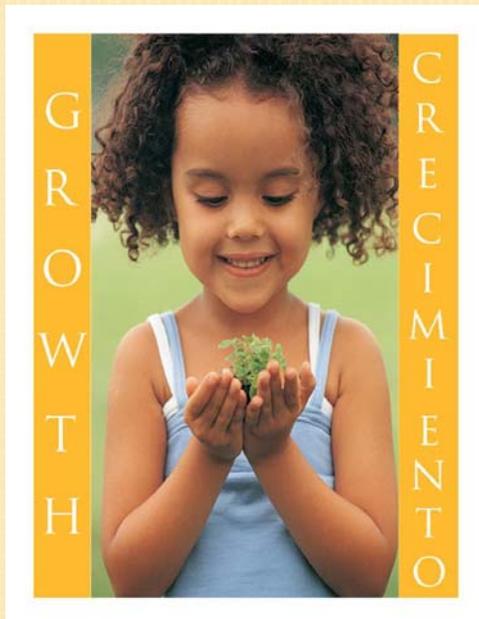
4. ENROLLING WOMEN WITH LATE OR NO PRENATAL CARE

- ✘ Enroll women based on their first prenatal care visit (regardless of timing)
- ✘ All women who did not receive prenatal care would be recruited from hospital

5. ENROLLING PRECONCEPTION COHORT

- ✘ Same sampling procedures for selecting non-pregnant as pregnant
- ✘ Additional recruiting sites (e.g., pregnancy testing clinics; household screening)

MAJOR QUESTIONS AND CHALLENGES OF PROPOSED APPROACH



Optimization of key parameters of provider-based sampling design

+ Balance:

- × a) # provider groups sampled with PSU;
- × b) # and % patients sampled within provider group;
- × c) impacts of decisions on geographic clustering;
- × d) need to minimize # hospitals
- × e) approach used to sample patients;

+ Affect design cost-efficiency and certain analytic objectives of NCS

+ Optimization depends on many factors: size of PSU, # and diversity of providers; scientific priorities

Potential bias from enrolling women from provider practices: late or no prenatal care

- + First trimester exposures linked to child health and development
- + 15-17% of women receive care after 1st trimester, and 1% receive no care prior to delivery
- + 1st trimester exposures likely different for women who receive early v. later care
- + Enrolling women regardless of timing of entry into prenatal care v. only those who receive care by 1st trimester

Sampling pregnant women within provider groups

- + Trade-offs between ease of operations, sampling efficiency and cost

Different sampling strategies for PSUs

- + PSUs vary considerably in provider profiles, ranging from 4 to >150 groups in our 10 counties
- + Optimal # of providers sampled may vary by county
- + Further discussion and guidance needed

Retaining geographic clustering in a provider-based sampling scheme

- + Retaining some level of geographic subsampling may be advantageous
- + May facilitate relatively efficient school-based data and environmental sample collections
- + Conversely, may not represent range of environmental and social exposures in PSU

6. Operational challenges to sampling provider groups

- + Having accurate sampling frame
- + Linking individual providers to practice groups
- + Finding accurate sources of practice volume

7. Stratification factors for sampling providers

- + Sampling providers based on provider v. patient characteristics
- + Obtaining accurate data on patient characteristics by provider

8. Type of providers to include in sampling frame

- + Include v. not include other types of providers (e.g., social service providers)
- + Potential for higher probability of being sampled if seen at both types of provider

9. Representing difficult-to-enroll subpopulations

- + Willingness to participate affects ability to represent births in county
- + If over-sample difficult groups, may need to deviate from PPS but in carefully controlled way

10. Challenges in obtaining a pre-conception cohort

- + Current NCS algorithm using pregnancy planning for frequency of follow-up results in biased cohort
- + Provider-based sampling is similarly biased (seekers of gynecological care)
- + If pre-conception remains priority, supplement strategies to identify sexually active women for followup

11. Non-participating provider groups

- + Replace refusals with provider group within same stratum
- + Review provider compensation and incentives

-
12. Ongoing changes to recruitment protocol over enrollment period
- + Use birth data to inform changes to protocol when deviations to representativeness noted
 - + Update sampling frame
13. Probability v. non-probability sampling
- + Probability sampling: known mathematical and statistical inferential properties, and >75 years of research to account for impact of practical implementation issues on inference
 - + Non-probability approaches: *can* enroll representative sample, and may enroll families who are more likely to be retained for 21-year follow-up

TAKE HOME MESSAGES

- ✘ Provider-based SAMPLING (vs. recruitment) has considerable potential merit.
- ✘ Three new Study Centers piloting this approach.
- ✘ Related efforts proposed and recommended for the current 10 sites.
- ✘ Strongly encourage increased interaction between these two pilot efforts.
- ✘ Need to refine targeted evaluation plan to assess cost and coverage (representativeness) of these approaches.

-
- ✘ For more details, please refer to the paper presented to the NCS: ¹Implementing Provider-Based Sampling for the National Children's Study: Opportunities and Challenges. A Draft Concept Paper Developed by the Ten Counties Participating in the "Provider-Based Recruitment Strategy" of the Vanguard Phase of the NCS. May 23, 2011.