



Discussion of potential sampling strategies for the National Children's Study, Main Study

May 29, 2012

*Moderator: Enrique Schisterman, Ph.D.
Senior Investigator
Division of Epidemiology Statistics and Prevention Research
Eunice Kennedy Shriver National Institute of Child Health
and Human Development, NIH*

Discussion: Geographic Area Sampling



- What should the probability sample look like?
 - Can we have a purely probabilistic geographic sample (no oversampling, weighting, or stratification)? What is gained or lost by this approach?
 - Should the geographic samples be clustered within regions? Would this clustering need to be defined by population density, or could it be defined by environmental characteristics? What is gained or lost by this approach?
 - Should the geographic samples be selected using a stratified frame (by population density, demographic characteristics, happiness index, etc.)? What is gained or lost by this approach?

Discussion: Geographic Area Sampling (*continued*)



- What should the probability sample look like?
 - How could frame deficiencies be identified, and how could they be backfilled?
 - How many geographic areas need to be selected in order to generalize the findings of the study?
 - What should the area of the sampled geographic units be (i.e. state, county, zip code, census tract, census block group, or census block)?

Discussion: Provider Sampling



- How can providers be enumerated efficiently?
- How can selected providers that choose not to engage in the study be replaced or substituted in a way that preserves the probability sample?
- How can the method of selecting providers increase recruitment success (i.e. restricted frame sampling)?
- Can the demographic characteristics of the provider's practice be determined prior to sampling?

Discussion: Provider Sampling (*continued*)



- Can the sampling method be flexible in this stage in order to allow for regional differences in identification of providers (such as availability of birth certificate records)?
- Are there features of a provider practice (practice type, size, etc.) that might bias the recruitment of participants?

Discussion: Sampling within Providers



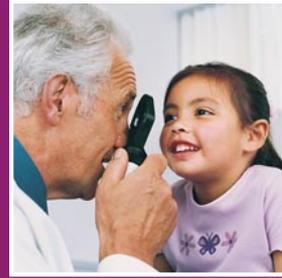
- What are effective ways to enroll women using systematic sampling? What could the basis of this sample be? Are there additional ways women could be sampled in an effective way that preserves the probability sample?
- Should women residing outside of the geographic sample, but seeking care from a selected provider, be included? If they are included, what would this do to the comparisons to extant natality or American Community Survey data?

Discussion: Sampling within Providers *(continued)*



- Should pre-conception women be eligible in the probability sample, or should they be from a separate cohort?
- How would the sample of pregnant women be evaluated for frame coverage, or population representation? How could deficiencies be addressed in this stage of sampling?
- How could women who change providers be handled?
- How could women who move out of the geographic area be retained?

Discussion: Sampling within Providers *(continued)*



- How can additional women, who reside within the geographic sample, but are not in the provider sample be included in a “light touch” cohort (for example, could a provider recruit women from a practice location other than the selected one) as a supplement?

Discussion: Supplementary Sampling



- Example 1: Women without prenatal care access; would it be adequate to recruit women from hospitals or birthing centers, excluding those who received prenatal care?
- Could a sample like this be recruited as a convenience sample?
- Would the women recruited in this way be considered a sub study, or could they be a part of the larger sample?
- When should a supplemental sample become a separate sample frame?

Discussion: Supplementary Sampling (*Example 1 cont.*)



- Could meta-analysis techniques be used to combine a supplemental frame with the larger probability-based cohort? For example, if a hypothesis was posed about left-handedness and an exposure, could the information be pooled from both cohorts with regards to the exposure-outcome relationship?
- What sources of bias would you anticipate by introducing the supplemental frame?

Discussion: Supplementary Sampling



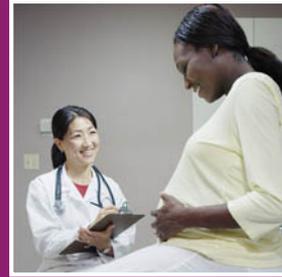
- Example 2: If lower income women are found to be deficient in the frame: could WIC providers or other list frames be added to the provider frame, with exclusion criteria for women who have already seen one of the selected providers?
- Could a sample like this be recruited as a convenience sample?
- Would the women recruited in this way be considered a sub study, or could they be a part of the larger sample?

Discussion: Supplementary Sampling (*Example 2 cont.*)



- When should a supplemental sample become a separate sample frame?
- Could meta-analysis techniques be used to combine a supplemental frame with the larger probability-based cohort? For example, if a hypothesis was posed about left-handedness and an exposure, could the information be pooled from both cohorts with regards to the exposure-outcome relationship?
- What sources of bias would you anticipate by introducing the supplemental frame?

Discussion: Missingness by Design



We are considering the use of a core questionnaire for everyone, with additional modules, or datasets for subsets of the study population.

- Can we have different questionnaire intensities within the larger frame?
- Can we have different questionnaire intensities between the supplemental frames and the probability sample?
- What would be the parameters for determining the sample sizes of women receiving low intensity and high intensity instruments?