ON DEVELOPING A NATIONAL SAMPLE OF WOMEN EARLY IN PREGNANCY

William D. Kalsbeek, Sharon Y. Cornell, Patricia S. Tennis
University of North Carolina at Chapel Hill

1. BACKGROUND

In the face of extensive improvements in medical technology and significant reductions in maternal, fetal, and neonatal mortality, complications of pregnancy, delivery and the newborn period persist. The extent of these morbidities in the U.S. population, however, is largely unknown, as is the relative distribution across geographic and demographic boundaries. While some of the correlates of the morbidities (e.g., low socioeconomic status, extremes of age and parity) are well-established, causal relationships between specific determinants and complications are not well understood.

Three landmark studies of the epidemiology of adverse pregnancy outcomes were undertaken in the late 1950's and early 1960's: the Kauai Pregnancy Study (Bierman et al, 1963; Bierman et al, 1965), the British Perinatal Survey (Butler and Bonham, 1963), and the U.S. Collaborative Perinatal Project (Niswander and Gordon, 1972). Though these studies greatly increased the current knowledge of the distribution and causation of perinatal problems, they shed little light on the determinants of specific pregnancy complications. In addition, the conclusions drawn from the investigations were not readily generalizable to the pregnant population in the United States.

Despite their shortcomings, the findings from these studies have guided medical care and programmatic interventions for more than 20 years. But, their usefulness has become increasingly limited as the passing years have witnessed improvements in prenatal, intrapartal and neonatal care. The demographics of the population have shifted significantly. The U.S. pregnant population is now younger, has lower parities, and experiences more out-of-wedlock deliveries than the populations studied in the late fifties and sixties. Furthermore, more recent studies based on smaller samples have suggested that factors previously unstudied or not fully studied may be related to perinatal morbidities (e.g., sexually transmitted diseases, certain health behaviors, stress, occupation).

A broader perspective of health events surrounding pregnancy in the United States has been gained through several national surveys conducted since the early 1960's by the National Center for Health Statistics (1986). The data in these surveys, collected retrospectively through samples of vital records (i.e., birth and fetal death certificates), were obtained by a mail, self-reporting mode of data collection. Recall problems for some items and relatively low response rates (78 percent for women and 69 percent for their providers in the 1980 National Natality and Fetal Mortality Surveys) limit this otherwise extremely useful view of the national experience.

Perceived problems with existing data sources but mainly an interest in more detailed data than are currently available led to the funding in 1985 of a design study by the Centers for Disease Control.

The general goal of this study was to investigate reasonable options for choosing a national probability sample of noninstitutionalized women early in pregnancy for the purpose of obtaining detailed and somewhat more prospective data on the following measures:

1. Scope of maternal complications and other morbidity during pregnancy and up to one month postpartum;
2. Applications of prenatal health care including medications as well as diagnostic and therapeutic procedures;
3. Frequency and timing of prenatal care;
4. Labor and delivery practices;
5. Maternal behavioral characteristics;
6. Neonatal infant complications and other morbidity; and
7. Neonatal health care.

The survey addressing these objectives will be called the Pregnancy/Childbirth Survey in the sequel.

Although a specific definition of eligibility for the Pregnancy/Childbirth Survey is partially dependent on the design, two general criteria were set forth at the beginning of this design study. First, the onset of pregnancy must have occurred within some well-defined study period (of reasonable length to minimize any seasonality in the data). Second, the woman must intend to carry the pregnancy to term since CDC elected to exclude abortions in the scope of the investigation.

2. MOTIVATION FOR DESIGN OPTIONS

Several things would make sampling difficult in the Pregnancy/Childbirth Survey. First, with an estimated four million eligible pregnancies per year in the U.S. and about 64 million women of childbearing age, the subset of women at some stage of pregnancy becomes a relatively rare segment of the total population. Second, there are no centralized and machine-readable national lists of pregnant women, thus implying that direct list sampling of this special population cannot be done and that a valid national sample can only be obtained through screening of the general population or through a sample of those with whom they come in contact during pregnancy. In regards to the latter, two sources seemed plausible: the health providers from whom they receive prenatal care and the serological laboratories who confirm their pregnancy or the presence of prenatal morbidity. Three design options were therefore evident:

1. General Population Screening - A national sample of households is contacted by telephone and eligible pregnant women identified through these contacts;
2. Provider Approach - A national sample of prenatal health care providers is selected and eligible pregnant women being treated by them are identified; and
3. Laboratory Approach - A national sample of laboratories which analyze prenatal blood

tests applied during pregnancy is selected and eligible women are identified through these laboratories.

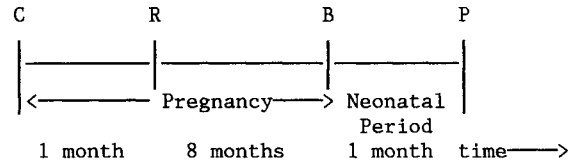
Before presenting a separate discussion of each design option, consider the following events occurring to a woman between onset of pregnancy and the end of the neonatal period after birth:

1. Conception;
2. Recognition of pregnancy;
3. Birth of the child, or children in the case of multiple births, and;
4. Postpartum, specifically 28 days after birth, at which point the neonatal period ends.

Figure A demonstrates how the above-mentioned events would appear in time sequence for an assumed "typical" full-term pregnancy, with the length of the time interval separating the events given in months.

FIGURE A

Major events during pregnancy and early postpartum care



To reduce the amount of recall error, a more prospective approach was thought to be preferable. A prospective design requires that the sample be limited to those women who have recently discovered that they are pregnant. This cohort of pregnant women would then be contacted periodically until the pregnancy is terminated, or until the end of the neonatal period. At each followup, members of the sample cohort would provide information retroactive to the last followup contact. The maximum duration of followup would therefore be around nine months, assuming that the length of time between events C and R is approximately one month. Finally, the cohort followed under the prospective approach would be "closed" in the sense that only those women sampled initially would be followed. Those moving out of the state would be followed (assuming adequate resources exist to do so) and no supplementary sample of eligible women moving into the state during the period of study would be selected.

Discussion of the design options we have considered will be presented as follows. Each design option will first be described and then briefly discussed. The format of discussion will be to simply list some of the strengths and potential weaknesses of each option. After the presentation of options, we will suggest some issues to consider in choosing among the options.

3. DESIGN OPTION 1: GENERAL POPULATION SCREENING

Under the first design option one would screen the general noninstitutionalized household population for women who found out they were pregnant within some established time period. For example, the general vehicle for population sampling might be through interviews applied to some sort of random digit telephone sampling design. Screening would be done by asking each woman in selected households whether she, and perhaps another woman within some well-defined social network, had discovered that she was pregnant (i.e., experienced event R) during the past M months.

Random digit telephone sampling could be accomplished through any of several mechanisms; e.g., Waksberg (1978) two-stage sampling, simple random sampling of numbers with a large percentage of nonresidential and nonworking numbers removed; and next digit selection of numbers from telephone directories. The choice among random digit sampling methods would depend on the type of telephone number frames that are available. In those instances where lists purged of nonworking and nonresidential numbers do not exist, the Waksberg selection procedure mentioned above might be used. Oversampling areas with relatively higher concentrations of eligible women (through stratification or PPS sampling) would further improve the efficiency of the screening process if eligible women tend to be geographically clustered (Kalton and Anderson, 1986).

Advantages

1. Overall response rates would be higher since women are contacted directly for participation in the study, instead of through third parties (e.g., physicians, hospitals, health departments), which add a source of attrition. In addition to the possibility that some providers will be unwilling to help in locating eligible women, sampling through participating providers raises the spectre of informed consent and other related logistical problems which increase the likelihood that a woman selected for participation will become a nonrespondent.
2. The sampling frame would be simple, namely one of those normally used for random digit telephone sampling.
3. The investigator can more precisely control when during pregnancy a woman will be contacted to participate in the study. The other approaches depend on when certain services or health care is received during pregnancy. This control allows one to weigh the cost-effectiveness of the length of time after event R that a woman would be considered eligible for the study. The longer the interval allowed after event R, the higher the percentage of women sampled who will be eligible but the greater the likelihood that certain health events might have already occurred, thus relying on the woman's recall for data.

Disadvantages

1. Because telephones do not serve all homes in the United States, there is a potential for a coverage problem. It is well known, for example, that although less than 10 percent of homes do not have telephones, those without telephones tend to be more rural, black and uneducated (Groves and Kahn, 1979). In areas where the coverage issue is especially acute, it may be

useful to supplement any telephone screening with a sample of nontelephone households to be screened in person. Coverage problems could also be partially adjusted for during analysis.

2. Women may be reluctant to divulge pregnancy information over the telephone, especially if the pregnancy is unwanted. The statistical implication of this potential problem would be a further coverage bias affecting especially those in whom there may be considerable interest (e.g., young unmarried women). This matter would be an issue with all three approaches to sampling, however.
3. While the telephone screening approach is simple, sound and "clean" statistically, perhaps its principal limitation is that it would be cumbersome and very expensive operationally. To examine the cost issue and ways for dealing with it, let us first consider a measure of operational efficiency (in fact, inefficiency), E_1, defined as the average number of telephone numbers that must be called (without network sampling being used) to locate a household with an eligible woman present. To compute E_1 for a national survey we assume that a household has no more than one eligible woman present at any time and that recognition (i.e., event R) is uniformly distributed through time. A rough measure of E_1 is then,

$$E_1 = \left(\frac{12H}{PGM}\right)$$

where H is the number of residential households, P is the number of pregnancies (excluding induced abortions) in a year's time, and G is the proportion residential among all telephone numbers that are called under a given sampling design. It has been estimated that H=80 million (U.S. Bureau of the Census, 1983) and P=4.0 million (U.S. Department of Health and Human Services, 1981, p. 201).

Values of E_1 for various assumed values of M and G are presented in Table 1. The data in this table reveal that even under an approach to random digit dialing where G is near one, a disturbingly large number of telephone numbers, on average, must be called to reach an eligible pregnant woman. For example, assuming one month between events C and R, about 12,000 telephone numbers would have to be called in order to locate 100 women who are in their first trimester of pregnancy. Approximately 4,800 telephone numbers would be needed if one wished to identify 100 women at anytime during her first two trimesters. Using a Waksberg approach to random digit dialing where G might be between 0.50 and 0.75, it would take between 6,400 and 9,600 telephone numbers to reach 100 women in their first or second trimester. It is clear from these figures that only in situations where M and G are both high would at-large telephone screening, without some effort to improve the calling success rate, be worth considering.

One way to improve one's chances of discovering eligible females during screening would be to sample at disproportionately higher rates in telephone clusters with higher concentrations of eligible women, although these may be somewhat difficult to identify and concentrations may not be high enough to make the effort worthwhile. Another way to increase the location rate for eligibles would be to extend the population screened through selected households to include a well-defined group of women living outside of those in selected households (e.g., women living in the households occupied by brothers and sisters of the head of household, women in households on the same block as the selected household, etc.).

Compared to the simple approach where screening is limited to members of selected households, this adaptation, called network sampling, would increase the number of identified eligibles in direct proportion to the average size (μ_k) of the network, which for purposes of discussion here includes the initial respondent, women in her household, and the members of the network outside of her household. The operational improvement in the screening process brought about by network sampling can be expressed in terms of a second type of telephoning efficiency (E_2) which measures the average number of eligible women identified per screening or interviewing call attempt. For a screening protocol where only one female of childbearing age per household is asked if she is pregnant, this efficiency can be measured as $E_2=\pi_M/C_I$ where π_M is the proportion of women of childbearing age in the population who have known for M months that they are pregnant, and C_I is the average number of call attempts made per initial respondent. For a comparable network protocol where the woman is asked if she or any member of her network is pregnant, the corresponding efficiency is $E_2=\pi_M\mu_k/C_I[1+r(\mu_k-1)\pi_M]$, where $r=C_N/C_I$ and C_N is the average number of calls made to interview eligible women identified as members of the network of an initial respondent. The efficiency of the network design relative to the non-network design is therefore,

$$\Phi=\mu_k/[1+r(\mu_k-1)\pi_M]. \qquad (1)$$

Assuming that the 63.6 million women of childbearing age account for the 4 million pregnancies per year, Table 2 presents values of Φ for different settings when the intent is to screen for pregnancy through the first two trimesters. As expected the increased efficiency of network sampling is of the order μ_k, although somewhat below this level as π_M, μ_k, and r increase.

Accompanying this increased operational

efficiency, however, would be a possible loss in statistical efficiency due to variable weighting brought about by the fact that the chances of discovering an eligible woman is directly proportional to the size of her network (Sirken, 1972). Thus, if the loss in statistical efficiency is directly related to μ_k, which in turn is inversely related to the loss in operational efficiency, then an optimum μ_k which jointly minimizes both losses must be found. Another consideration in choosing an appropriate network is the inverse relationship between the accuracy of reporting eligibles in the network and μ_k for some kinds of rare population attributes (Sudman, 1985).

4. DESIGN OPTION 2: PROVIDER APPROACH

It is recommended that prenatal care start as early in the first trimester as possible. For uncomplicated pregnancies, standard prenatal care should involve visits every four weeks during the first 28 weeks of pregnancy, once every two weeks for the next eight weeks, and once a week until delivery (AAP/ACOG, 1983). During an average pregnancy a woman would visit her prenatal care provider(s) approximately 10-12 times according to these recommendations. For most pregnant women, these recommendations are met. According to data from the National Center for Health Statistics (1984a), 74.2 percent of all women giving birth in 1982 started prenatal care during the first three months of pregnancy.

Because the majority of pregnant women start their prenatal care during the first trimester, one could therefore obtain a sample of pregnant women early in pregnancy by sampling prenatal care providers. Within selected providers a sample of eligible pregnant women would be chosen, perhaps using time intervals as secondary sampling units. Prenatal care "providers" include individual professionals such as physicians, nurse practitioners and midwives, who provide care individually as part of health care organizations such as office-based physician practices, hospitals, and public or private health clinics. Sampling frames for these providers may come from any of several (possibly overlapping) sources, including state licensure boards and membership listings for various professional organizations as well as societies dealing with maternal health care.

Advantages

1. A high percentage of women eventually see a provider for prenatal care, thus making good sample coverage possible.
2. If women must be contacted through secondary sources, the provider will be the most likely person to know the woman's address, thus increasing the likelihood that a selected subject would be contacted.

Disadvantages

1. The provider frames may be incomplete or outdated thus causing biases in estimates due to miscoverage.
2. Because information to identify and locate selected women would be obtained through a health care provider and since many providers (with a record of low survey response rates) might be reluctant to provide such information, response rates by this sampling option may be relatively low.
3. Using multiple overlapping sampling frames to select the provider sample requires compensatory measures similar in principle to those with network sampling, but potentially much more complex.
4. There could also be major operational problems in piecing together a national provider frame consisting of lists of varying quality and content from several national organizations and societies.

Problems with sampling from multiple frames motivates us to seek a simpler solution to provider sampling. To find one, consider the following on the usual locations of prenatal care:

Location of Care	Approx. % of Pregnancies	Physician Involved?
1. Private Physician	55	Yes
2. Health Department	15	Yes
3. Health Maintenance Organization	10	Yes
4. Uniersity Hospital Clinics	8	Yes
5. Military Facilities	<5	Yes
6. Federally Funded Community Health Ctrs.	<5	Yes
7. Migrant Health Centers	<5	Yes?
8. Indian Health Service	<5	Yes
9. Community Based Hospitals	<5	Yes

Noting from the above table that physicians seem to be involved at some point during prenatal care, consider the feasibility of sampling pregnant women through physicians alone. The logic for this idea is further supported by data from the National Medical Care Utilization and Expenditure Survey (NMCUES), where we found that all 52 of the full-term pregnant respondents eventually saw a physician or were treated by someone under the supervision of a physician. The availability of a single high quality frame for physicians through the Record of Physician Activities and the American Medical Association makes this version of the provider approach all the more attractive (AMA, 1979).

This physicians-only adaptation of the provider approach is not without its problems. First, physician involvement in providing prenatal care to some women may not be direct, in the sense that the physician provides the care in person. For example, women in public or private community health centers may only see nurse practitioners or physicians assistants exclusively and not the physician who oversees the care given and never actually sees the woman during her visits. This distancing of the physician from some women may increase the likelihood that they would be overlooked during the case ascertainment phase of the survey. Second, women who first receive care from a nonphysician provider will begin their participation later in pregnancy than they would have under the provider approach where all provider types were sampled. This later start-up for some women would increase the length of time for

which they must recall previous events in their pregnancy, thus contributing further to nonsampling error.

5. DESIGN OPTION 3: LABORATORY APPROACH

Once a woman finds out she is pregnant and decides to carry the pregnancy to term, she usually sees a health care provider for her first prenatal exam. It is highly recommended (AAP/ACOG, 1983) and legislated in most states that a blood sample be drawn during this exam in order to test for presence of syphilis and exposure to rubella. Because of their role in this process, the laboratories which perform these serological tests could be used as a basis for sampling pregnant women.

The protocol under this laboratory approach to sampling would involve at least two stages. In the first stage laboratories, possibly stratified by type of laboratory might be selected. The second stage could involve sampling women within each sampled laboratory and would require compiling a list of patients receiving prenatal blood tests. The women sampled from the laboratories' patient lists must then be contacted and recruited into the study.

To assess the difficulty of forming a frame for sampling laboratories, all professional organizations to which the personnel of laboratories or the laboratories themselves subscribe as well as agencies involved in licensing or regulating the laboratories were identified. No single source provided an adequate national frame for these laboratories. Frame construction would thus require merging a relatively large number of partially complete and overlapping lists. An added complication in forming the list was that not all these lists are accessible or in the same units (i.e., laboratories versus professional names).

Advantages

1. Testing in these laboratories is often done soon after event R in a woman's pregnancy period, thus enabling study subjects to be followed prospectively through much of this period.
2. Once the PSU frame is constructed the sampling design within stages would be relatively simple provided that things such as multiplicity (caused by women being tested more than once) did not occur.

Disadvantages

1. The coverage rate in a study using this design option would be partly dependent on the percentage of women who are tested by the laboratories included in the sampling frame. In states where testing for venereal disease is not required, the approach would probably be infeasible since coverage would be too low.
2. There is a potential for considerable difficulty in identifying all laboratories performing these tests. Since there is often no professional linkage among them (e.g., through membership in a professional organizaation, licensure boards, etc.), the list of laboratories would have to be pieced together from several quite diverse sources (e.g., Health Care Financing Administration, American Association of Clinical Pathologists, National Association of Counties, Medical Laboratory Observer subscriber list).
3. Overall response rates for women might be quite low since contact for recruiting women for the study would require at least one intermediate contact, namely the provider who drew the blood and requested the analysis. The woman whose blood was tested might also need to sign some kind of informed consent form before attempts to contact her can be made.

6. DISCUSSION

We conclude by briefly discussing some things to consider in choosing among the options presented above:

1. The specific nature of data to be collected must be considered since, for example, requiring large amounts of technical medical information would make gaining access to provider records more important and thus make sampling providers a more attractive approach.
2. The availability of frame information (i.e., for health care providers, serological laboratories and telephone numbers) will help to determine which of the options is the most realistic from a sampling standpoint. For example, severe problems in dealing with the multiple frame sources under the provider and laboratory approaches might make general population screening most practical.
3. The expertise of personnel available to conduct the study would partially determine the feasibility of various options. If the study were to be conducted by those experienced in doing provider surveys and sampling provider medical records, then any sampling option would be feasible. When those conducting the survey have little prior experience, telephone screening option may very well be the most realistic.
4. The amount of measurement error likely might also affect the choice among options. Assuming that shorter recall periods for measures of morbidity and health care are desirable, a slight advantage among options would be with the screening portion, where women could potentially be enrolled in the study sooner after recognition than the other approaches. The length of recall for prenatal events, assuming that women are interviewed soon after their first prenatal provider visits, would be distributed approximately as follows (NCHS, 1984a):

Length of Recall (In Months)	Percent of Pregnancies
1-2	50
3	25
4-6	20
≥ 7	5

REFERENCES

American Academy of Pediatrics/American College of Obstetricians and Gynecologists, Guidelines for Perinatal Care, (Evanston, IL: American Academy of Pediatrics/American College of Obstetricians and Gynecologists), 1983.

American Medical Association, Division of Survey and Data Resources, The Physician Masterfile, 1979.

Bierman JH, Siegel E, French F, Connor A, "The Community Impact of Handicaps of Prenatal or Natal Origin," Public Health Reports, 78: 839, 1963.

Bierman JM, Siegel E, French F, Simonian K, "Analysis of the Outcome of All Pregnancies in a Community," American Journal of Obstetrics and Gynecology, 1: 37, 1965.

Butler NR, Bonham DG, Perinatal Mortality: The First Report of the 1958 British Perinatal Mortality Survey, (London: E & A Livingston, Ltd.), 1963.

Groves RM, Kahn RL, Surveys by Telephone, (New York: Academic Press), 1979.

Kalton G, Anderson DW, "Sampling Rare Populations," Journal of the Royal Statistical Society, Series A, 149(1): 65-82, 1986.

National Center for Health Statistics (1978), "Wanted and Unwanted Childbearing in the United States: 1968, 1969 and 1972 National Natality Surveys," Vital and Health Statistics, Series 21, No. 32, U.S. Government Printing Office, Washington, D.C.

National Center for Health Statistics (1984a), Monthly Vital Statistics Report, Advance Report on Final Natality Statistics, 1982, vol. 33, No. 6, supplement, September 1984.

National Center for Health Statistics (1984b), "Patterns of Ambulatory Care in OB-GYN," Vital and Health Statistics, January 1980-December 1981, Series 13, No. 76, February 1984.

National Center for Health Statistics (1986), "Methods and Response Characteristics: 1980 National Natality and Fetal Mortality Surveys," Vital and Health Statistics, Series 2, No. 100, January 1986.

Niswander KR, Gordon M, The Collaborative Perinatal Study of the National Institute of Neurological Diseases and Stroke: The Women and Their Pregnancies, (Philadelphia: WS Saunders), 1972.

Sirken MG, "Stratified Sample Surveys with Multiplicity," Journal of the American Statistical Association, 67: 224-227, 1972.

Sudman S, Kalton, G, "New Developments in the Sampling of Special Populations," Annual Review of Sociology, 12: 410-429, 1986.

Sudman S, "Experiments in the Measurement of the Size of Social Networks," Social Networks, 7: 127-151, 1985.

U.S. Department of Health and Human Services, Better Health for our Children: A National Strategy, Report of the Select Panel for Promotion of Child Health, Vol. 4, Background Papers, USGPO, Washington, D.C., 1981.

U.S. Bureau of the Census, "Population Profile of the United States: 1982," Current Population Reports, Series P-23, No. 130, Special Studies, USGOP, Washington, D.C., 1983.

U.S. Bureau of the Census, "Estimates of the Population of the United States, by Age, Sex and Race: 1980 to 1986," Current Population Reports, Series P-25, No. 1000, USGPO, Washington, D.C., February 1987.

Waksberg J, "Sampling Methods for Random Digit Dialing," Journal of the American Statistical Association, 74: 40-46, 1978.

TABLE 1. Telephone Efficiency (E_1) for At-Large Population Screening for Women Who Have Discovered Within the Past M Months That They are Pregnant*

Number of Months Since Discovery (M)	Proportion of Telephone Numbers Called That Would Reach a Residence** (G)			
	0.25	0.50	0.75	1.00
1	960	480	320	240
2	480	240	112	120
3	320	112	107	80
4	240	120	80	60
5	192	96	64	48

*"Telephone efficiency" is measured as the average number of telephone numbers one must call in order to reach a household with a woman who has discovered within the past M months that she is pregnant.

**The proportion (G) would vary depending on the distribution and extent of residential telephones in the population being covered and the approach to random digit dialing that one followed.

TABLE 2. Increased Efficiency of Network Sampling (Φ) in At-Large Telephone Screening of Women Discovering M Months Ago That They are Pregnant*

Number of Months Since Discovery (M)	Average Size of Network (μ_k)	Relative Average Number of Call Attempts: Initial Respondents Versus Network Members (r)		
		1.0	1.5	2.0
First Trimester:**				
2	2	1.98	1.97	1.96
(π_M)=0.0105	6	5.74	5.61	5.49
	10	9.24	8.90	8.58
Second Trimester:**				
5	2	1.96	1.93	1.91
(π_M)=0.0262	6	5.38	5.12	4.88
	10	8.29	7.64	7.08

*Computed as $\Phi=\mu_k/[1+r(\mu_k-1)\pi_M$, where

$$\pi_M = \frac{[4 \text{ million pregnancies}][M/12]}{[63.6 \text{ million women of CB age}]}$$

**Presumes one month duration between conception (event C) and recognition (event R).