Source of Bias in Prenatal Care Utilization Indices: Implications for Evaluating the Medicaid Expansion

Greg R. Alexander, MPH, ScD, Mark E. Tompkins, PhD, Donna J. Petersen, MHS, ScD, and Judith Weiss

ABSTRACT

Background: Recent expansions in eligibility for coverage of prenatal care services by the Medicaid program reflect national initiatives to improve pregnancy outcomes. This study investigates the potential impact that completeness of reporting of prenatal care and gestational age variables and strategies to impute missing data may have on evaluations of the Medicaid expansion.

Methods: This study, examining 15 years of vital record data from a single state and comparing 1 year of data from four mid-Atlantic states, selected single live births to resident mothers for analyses. The "day 15" and the "preceding case" methods were used to impute missing gestational age data.

Results: Considerable temporal and geographic variation was detected in completeness of reporting of variables used to construct prenatal care indices. After imputing values for cases with missing data, the proportion of cases for which adequacy of prenatal care utilization could not be determined ranged from 3% to 24% among the states investigated. For those cases where gestational age data could be imputed, the distribution of prenatal care utilization was not markedly disparate from those cases with complete reporting of gestational age.

Conclusions: The results indicate that variations in reporting, decisions regarding the treatment of missing data, and the choice of the denominator can alter prenatal care utilization percentages and have implications for evaluations of the impact of the recent Medicaid expansion on prenatal care utilization. (*Am J Public Health.* 1991;81:1013–1016)

Introduction

Recent expansions in eligibility for coverage of prenatal care services by the Medicaid program reflect policy initiatives to meet Year 2000 national health objectives related to low birthweight and infant mortality.[1] Increased utilization of prenatal care has been associated with improved pregnancy outcomes.[2–7] In an effort to reduce financial barriers to prenatal care, Medicaid eligibility criteria were raised in 1990 to 133% of the federal poverty level, with state options to extend coverage to 185% of poverty. One focus for evaluation of this policy is the extent to which changes in eligibility for state Medicaid coverage will increase utilization of prenatal care and ultimately decrease low birthweight percentages and infant mortality rates.[8]

Prenatal care utilization is commonly measured by composite indices, the most notable being the Kessner Index of Adequacy of Prenatal Care.[2] This index considers the month prenatal care began, the number of prenatal care visits received, and the gestational age at delivery to categorize prenatal care utilization as adequate, intermediate, or inadequate. The adequate category includes cases with a first trimester initiation of care and prenatal care visits that are more than or equal to a recommended number for a pregnancy of a given duration. More visits are required as gestation increases. Initiation of care in the second trimester or an earlier initiation with less than the number of visits recommended to achieve the adequate category is termed intermediate care. Cases with no prenatal care, initiation of care in the third trimester, or an earlier initiation of care with less than the number of visits stipulated to achieve the intermediate category are labeled inadequate care. Modifications and alternatives to this index have been proposed[3,9–11] and used to investigate the relationship of prenatal care to a variety of pregnancy outcome measures.[2–6,9] Unfortunately, descriptions of these indices have often failed to specify how missing or implausible data should be handled.

The treatment of missing data in the construction of these indices is an important issue. In the Kessner index, cases with missing prenatal care data are classified as having inadequate prenatal care utilization.[2] Further, while various methodologies have been suggested to impute missing gestational age data,[12–15] a convention for the use of imputed gestational age data in the calculation of prenatal care utilization indices has not been established. The date of last menstrual period (LMP), needed to calculate the gestational age interval, has been estimated to be incomplete on approximately 20% of the nation's birth certificates.[12]

State-specific analyses of trends in prenatal care utilization before and after adoption of Medicaid options and comparisons of prenatal care among states with varying eligibility levels may be sub-

Greg R. Alexander is with the Maternal and Child Health Program, School of Public Health, University of Minnesota. Mark E. Tompkins is with the Department of Government and International Studies, University of South Carolina. Donna J. Petersen is with the Department of Health, State of Minnesota. Judith Weiss is with the Department of Health Policy and Management, The Johns Hopkins University.

Requests for reprints should be sent to Dr. Greg R. Alexander, Maternal and Child Health, University of Minnesota, PO Box 197 Mayo, 420 Delaware Street, SE, Minneapolis, MN 55455.

This paper was submitted to the journal September 27, 1990, and accepted with revisions March 6, 1991.

TABLE 1—Percentage of Missing Components of Prenatal Care Utilization Index by 5-Year Groups: 1974–1988 South Carolina Resident Single Live Births

Missing Component	1974–1978	1979–1983	1984–1988
		Percent	
Trimester that prenatal care began[a]	5.99	0.74	1.10
No. of prenatal care visits[b]	1.09	0.62	1.04
Gestational age interval[c]			
No imputation	16.33	5.31	6.91
Day 15	2.09	1.97	2.21
Preceding case	2.04	1.92	2.15
Total missing[d]			
No imputation	19.05	5.37	7.03
Day 15	7.50	2.21	2.63
Preceding case	7.47	2.18	2.58

[a]Includes cases with a not-stated month prenatal care began or with a stated but nonspecific response, e.g., an actual month (January through December).
[b]Includes cases with a not-stated number of visits or with stated visits of 50 or more.
[c]Includes cases with an incomplete date of last menstrual period or a missing gestational age interval value after imputation, where indicated. Gestational age intervals <20 or >50 weeks are treated as missing.
[d]Includes cases with a missing value for any of the prenatal care utilization index component variables, i.e., trimester prenatal care began, number of prenatal care visits, and gestational age interval from last menstrual period.

TABLE 2—Prenatal Care Utilization Percentages by Gestational Age Imputation Method and 5-Year Groups: 1974–1988 South Carolina Resident Single Live Births

Prenatal Care Groups	1974–1978 (N = 237 347)	1979–1983 (N = 251 247)	1984–1988 (N = 256 175)
No imputation			
Adequate	48.36	55.81	53.61
Intermediate	24.43	29.23	28.45
Inadequate/missing	27.21	14.97	17.94
Missing[a]	19.05	5.37	7.03
Inadequate	8.17	9.59	10.91
Day 15			
Adequate	54.27	57.42	55.88
Intermediate	28.46	30.34	29.96
Inadequate/missing	17.27	12.23	14.16
Missing[a]	7.50	2.21	2.63
Inadequate	9.77	10.02	11.53
Preceding case			
Adequate	54.27	57.43	55.89
Intermediate	28.48	30.36	29.98
Inadequate/missing	17.25	12.21	14.13
Missing[a]	7.47	2.18	2.58
Inadequate	−9.78	−10.03	−11.55

[a]Includes cases with a missing value for any of the prenatal care utilization index component variables, i.e., trimester prenatal care began, number of prenatal care visits, and gestational age (after imputation where indicated).
Note: Because of rounding, columns may not total 100%.

ject to bias due to changes in the completeness of reporting of the variables used to construct these indices and in methodologies employed to impute missing data. This study investigates the potential impact that completeness of reporting of prenatal care and gestational age variables and strategies to impute missing data may have on evaluations of the Medicaid expansion. Two study designs have been employed: (1) examining 15 years of data from a single state, and (2) comparing 1 year of data from four eastern states.

Data and Methods

Computerized South Carolina live birth files for the years 1974 to 1988 were used for the first component of this investigation. For the second component, additional data from three mid-Atlantic states with varying levels of completeness of reporting of gestational age were drawn from the Department of Health and Human Services (DHHS) Region III Perinatal Information Consortium's 1987 Live Birth File.[16] Single births to resident mothers were selected for both analyses.

Gestational age was calculated as the interval from the date of last normal menses to the date of birth and was truncated into completed weeks.[17] Two strategies were employed for imputing missing gestational age data.[15] The "day 15" method imputes a value of 15 for the day of last normal menses for those births for which this single piece of information prevents the calculation of the gestational age interval. The "preceding case" method imputes gestational age by inserting the calculated gestational age interval value of a preceding birth with a corresponding month of LMP, 500-g birthweight interval, and race of mother.

Results

Table 1 displays the proportion of South Carolina cases with missing prenatal care and gestational age information for three 5-year periods. Nearly 6% of the cases in the 1974 to 1978 period exhibited missing or nonspecific data on month or trimester of pregnancy that prenatal care began. Missing gestational age data were found for more than 16% of the cases in the 1974 to 1978 period, but dropped markedly in later periods. After imputation methods were employed to recapture cases with incomplete dates of LMP, gestational age was still missing for approximately 2% of the cases in each period.

Table 2 presents percentages of prenatal care utilization employing three different treatments of missing gestational age data (no imputation, day 15, and preceding case). Based on no imputation, adequate prenatal care utilization was lowest in the 1974 to 1978 period (48.36%). When cases with missing data were combined with the inadequate category, the percentages based on no imputation varied across the three periods from 27.21% to 14.97% to 17.94%. The separate inadequate category increased from 8.17% in the 1974 to 1978 period to 10.91% for 1984 to 1988. There is little appreciable difference in the prenatal care distribution percentages produced by the two alternative methods for imputing missing gestational age values.

The next phase of the analysis utilized 1987 data from four states. The preceding case method was selected to im-

pute missing gestational age values. Table 3 depicts by state the percentages of cases with missing gestational age data, before and after imputation, and missing prenatal care initiation and visit data. States B and C revealed a relatively higher proportion of missing gestational age data, 20.10% and 17.58%, respectively. While imputation markedly reduced this proportion in state C (5.73%), it had little impact in state B, where most cases with missing gestational age data were missing the entire date of LMP, not just the day value.

Prenatal care utilization percentages by state and gestational age imputation group are displayed in Table 4. These percentages are calculated in two different ways. In part A, percentages are based on a denominator of all live births with a separate category for cases with missing prenatal care or gestational age values after the application of the indicated gestational age imputation method. In part B, after imputation where indicated, cases with missing prenatal care and gestational age data are excluded, and percentages are based on a denominator of cases with the prerequisite in-range data elements needed to compute the index.

For states A and B, the imputation of gestational age had little appreciable effect on prenatal care percentages based on all cases (Table 4, part A). However, the proportion of total missing cases was noticeably reduced by imputing gestational age in the other two states. When calculations only included cases with in-range or imputed data values (Table 4, part B), the use of the gestational age imputation method had little appreciable impact on the prenatal care utilization percentages. This result suggests that the prenatal care utilization distribution for those cases for which a missing gestational age value could be imputed was only slightly less adequate than that of cases with complete gestational age data. For example, in state C, where gestational age was imputed for more than 10% of the cases, the adequate percentage was 71.55% before imputation and 71.06% afterward.

TABLE 3—Percentage of Missing Components of Prenatal Care Utilization Index by State: 1987 Resident Single Live Births

Missing Component	State A	State B	State C	SC[a]
Trimester that prenatal care began[b]	4.96	4.73	2.84	0.88
No. of prenatal care visits[c]	2.12	6.68	2.94	1.17
Gestational age interval[d]				
No imputation	5.10	20.10	17.58	7.26
Preceding case	4.69	20.01	5.73	2.16

[a]SC = South Carolina
[b]Missing trimester in pregnancy prenatal care began category includes cases with a not-stated month prenatal care began.
[c]Missing number of prenatal care visits category includes cases with a not-stated number of visits or with stated visits of 50 or greater.
[d]Missing gestational age category includes cases with an incomplete date of last menstrual period or a missing gestational age interval value after imputation, where indicated. Gestational age intervals of <20 or >50 weeks are treated as missing.

TABLE 4—Prenatal Care Utilization Percentages by Imputation Method and State: 1987 Resident Single Live Births

	State A	State B	State C	SC[a]
Part A: Percentages Based on All Cases				
No imputation	9929	70 835	158 468	51 580
Adequate	46.40	57.17	58.52	52.49
Intermediate	32.51	14.71	17.23	28.71
Inadequate	14.57	3.99	6.03	11.50
Missing[a]	6.52	24.13	18.21	7.30
Preceding case	9929	70 835	158 468	51 580
Adequate	46.57	57.22	65.90	55.02
Intermediate	32.68	14.73	19.87	30.29
Inadequate	14.60	3.99	6.97	12.14
Missing[a]	6.14	24.06	7.25	2.55
Part B: Percentages Based on Cases with a Calculated Prenatal Care Index Value (Missing Excluded)				
No imputation	9282	53 740	129 617	47 813
Adequate	49.63	75.36	71.55	56.62
Intermediate	34.78	19.39	21.06	30.97
Inadequate	15.59	5.25	7.39	12.41
Preceding case	9319	53 894	146 974	50 263
Adequate	49.62	75.34	71.06	56.46
Intermediate	34.82	19.40	21.42	31.08
Inadequate	15.56	5.26	7.52	12.46

[a]SC = South Carolina
[b]Includes cases with a missing value for trimester prenatal care began, number of prenatal care visits, or gestational age (after imputation where indicated).
Note: Because of rounding, columns may not total 100%.

Discussion

These data indicate that variations in reporting, decisions regarding the treatment of missing data, and choice of denominator can alter prenatal care utilization percentages. The true potential impact of the Medicaid expansion on prenatal care utilization may be obscured due to these factors. Therefore, evaluations of the Medicaid expansion should include a thorough examination of reporting completeness, a disclosure of strategies employed for imputing missing data, and a description of the methods used to calculate prenatal care percentages.

These data further suggest that it is unwise to categorize cases with missing prenatal care or gestational age data in the inadequate care group. The distribution of prenatal care utilization was not markedly disparate for those cases with imputed gestational age data and those cases with complete reporting of gestational age. This result argues both for employing gestational age imputation strategies and not assuming that cases with missing data had inadequate prenatal care utilization. While the prenatal care utilization of those cases with missing data remains unknown, estimates of inadequate prenatal care usage that include cases with missing data may be inflated.

Whether prenatal care percentages should be calculated from all cases or only

those for which the index can be generated is illustrated by data from state B. Based on all cases, 57% of women delivering in that state were placed in the adequate category. Excluding cases with missing data from the calculation resulted in a placement of 75% of women into the adequate category. The ease with which prenatal care percentages can be manipulated into adequate and other than adequate categories should not be ignored, given the political context of an impending evaluation of a major public policy initiative. If we assume that cases for which the index can be calculated are roughly similar or only marginally better in their prenatal care utilization patterns than those without the prerequisite data, the 75% figure may be a better estimate of the proportion of all births in the state with adequate prenatal care utilization. To alternatively assume that most cases with missing data have less than adequate utilization of prenatal care is unrealistic in areas with large proportions of missing data. While cases with missing gestational age data have been shown to display higher risks of poor pregnancy outcomes,[15] all cases with missing data will not necessarily exhibit adverse outcomes or less than adequate prenatal care utilization. Care should be taken in the analysis and reporting of these percentages to ensure that large missing group percentages are not used to represent individuals with less than adequate prenatal care utilization.

The choice between the day 15 or the preceding case method to impute missing gestational age data appears to have little impact on prenatal care utilization distributions. While the day 15 method has been noted to inflate preterm percentages in comparison with the preceding case method,[12,15] the amount of bias introduced by this approach appears insufficient to produce a prenatal care utilization distribution markedly different from that obtained by the preceding case method.

Prenatal care utilization indices are a useful but crude surveillance tool for assessing needs and evaluating programs and policy. Evaluation of the Medicaid expansion by these indices is a heuristic attempt to provide accountability for a major policy initiative that may enhance maternal and infant health in ways not easily appraised by available health measurement tools. Because the components of these indices are subject to reporting bias, and because important characteristics of prenatal care (e.g., quality, comprehensiveness, and continuity) are not considered, it is imperative that these indices be prudently employed in evaluation efforts and be continually scrutinized for bias.

Credible data for programmatic efforts and policy discourse require both accurate and complete reporting and logical and systematic analysis. The three component variables of prenatal care utilization indices contain potentially large measurement error. Vagaries in analytical methods used to compute these indices can further add to the uncertainty regarding the proposed relation between prenatal care utilization and pregnancy outcome and the possible impact of policy initiatives. These methodological concerns require ongoing attention, but they should not be used to divert attention from critically needed efforts to ensure that all women have access to and are encouraged to use prenatal care. □

Acknowledgment

This manuscript was presented in part at the 1990 annual meeting of the American Public Health Association in New York City.

References

1. From the Health Resources and Services Administration. Office of Maternal and Child Health. *JAMA.* 1989;261:199. Editorial.
2. Kessner DM, Singer J, Kalk CW, et al. Infant death: an analysis of maternal risk and health care. *In: Contrasts in Health Status*, I. Washington, DC: Institute of Medicine, National Academy of Sciences, 1973.
3. Gortmaker SL. The effects of prenatal care upon the health of the newborn. *Am J Public Health.* 1979;69:653–660.
4. Showstack JA, Budetti PO, Minkler D. Factors associated with birthweight: an exploration of the roles of prenatal care and length of gestation. *Am J Public Health.* 1984;74:1003–1008.
5. Alexander GR, Cornely DA. Racial disparities in pregnancy outcomes: the role of prenatal care utilization and maternal risk factors. *Am J Prev Med.* 1987;3:254–261.
6. Murray JL, Bernfield M. The differential effect of prenatal care on the incidence of low birth weight among whites and blacks in a prepaid health care plan. *N Engl J Med* 1988;319:1385–1391.
7. Institute of Medicine. *Preventing Low Birth Weight.* Washington, DC: National Academy Press, 1985.
8. Hogue C. Conference opening remarks presented at the Medicaid Expansion and Infant Mortality Workshop; June 1989; Bethesda, Md.
9. Alexander GR, Cornely DA. Prenatal care utilization: its measurement and relationship to pregnancy outcome. *Am J Prev Med.* 1987;3:243–253.
10. McDonald TP, Coburn AF. Predictors of prenatal care utilization. *Soc Sci Med.* 1988;27:167–172.
11. Kotelchuck M. The mis-measurement of prenatal care adequacy in the US and a proposed alternative two-part index. Presented at the APHA Annual Meeting; October 1987; New Orleans, La.
12. Taffel S, Johnson D, Heiser R. A method of imputing length of gestation on birth certificates. *Vital Health Stat.* [2]. 1982; No. 92.
13. Williams RL, Creasy RK, Cunningham GC, et al. Fetal growth and perinatal variability in California. *Obstet Gynecol.* 1982;59:624–632.
14. David RJ. Population-based intrauterine growth curves from computerized birth certificates. *South Med J.* 1983;76:1401–1406.
15. Alexander GR, Tompkins ME, Cornely DA. Gestational age reporting and preterm delivery. *Public Health Rep.* 1990;105:267–275.
16. Zinzeleta E: DHHS region III perinatal information consortium. Baltimore, Md: Johns Hopkins University; 1989. Technical Report 89-1.
17. Silverman WA, Lucey JF, Beard A, et al. Nomenclature for duration of pregnancy, birth weight and intrauterine growth. *Pediatrics.* 1967;39:935–939.