Maternal and Child Health Journal, Vol. 9, No. 2, June 2005 (© 2005)
DOI: 10.1007/s10995-005-4869-3

Clustering of U.S. Women Receiving No Prenatal Care: Differences in Pregnancy Outcomes and Implications for Targeting Interventions

Cathy R. Taylor, MSN, RN, DrPH,[1,3] Greg R. Alexander, RS, MPH, ScD,[2] and Joseph T. Hepworth, PhD[1]

Objectives: Prenatal care is an established mechanism for identifying and managing risk factors impacting pregnancy outcomes. Despite aggressive efforts in the United States (US) to assure that all women begin care in the first trimester, every year about 70,000 women in the US receive no care prior to delivery. We hypothesized that US women receiving no prenatal care comprise clusters (subgroups) with distinctive behavioral, socio-demographic, and medical risks and that birth outcomes differ among the clusters. *Methods*: White, Black, and Hispanic women ($n = 126,220$) receiving no prenatal care and delivering a live, singleton infant were identified from linked birth and death certificates for years 1995 through 1997. Cluster analysis was used to group women with similar characteristics, and cluster assignment was evaluated using discriminant analysis. Birth outcomes for any care and no-care women were then examined using logistic regression. *Results*: Six replicable clusters of women with no care were identified. Birth outcomes varied significantly among clusters and were two to four times worse for no-care clusters compared to outcomes for women receiving any care. *Conclusions*: Cluster analysis is an effective alternative for grouping individuals for use in public health education, intervention, and outreach programming. Women receiving no prenatal care were characteristically different from women receiving any care in this study, but they did not represent a homogenous group. Findings suggest that interventions should target reducing the proportion of women receiving no care and should be tailored to specific no-care clusters.

KEY WORDS: prenatal care; low-birth weight; preterm birth; infant mortality; cluster analysis.

Prenatal care is an established mechanism for identifying and managing medical, socio-demographic, and behavioral risk factors that may contribute to poor pregnancy outcomes. Despite longstanding and widespread federal and state efforts to assure that all women in the United States begin prenatal care in the first trimester, every year 1.5 to 2% of pregnant women (about 70,000) receive no care prior to delivery (1). No prenatal care represents a serious public health concern, as failure to enter care precludes screening and treatment for manageable conditions and essentially renders the pregnant woman invisible to the healthcare system.

Traditionally, studies have combined women receiving no care with women who enter care late or with those who receive less than the recommended number of prenatal care visits (2–4). Whereas the Kessner or Institute of Medicine (IOM) prenatal care utilization indices combined no care and inadequate care into one group, other care utilization indices, e.g., the GINDEX, proposed and recent analyses support separating these groups because of their

[1]School of Nursing, Vanderbilt University, Nashville, Tennessee.
[2]Department of Maternal and Child Health, University of Alabama at Birmingham, Birmingham, Alabama.
[3]Correspondence should be addressed to Cathy R. Taylor, MSN, RN, DrPH, Vanderbilt University School of Nursing, c/o Meharry-Vanderbilt Alliance, 1919 Charotte Avenue, Suite 300, Nashville, Tennessee 37203; e-mail: cathy.taylor@ vanderbilt.edu.

different birth outcomes (2–9). Not using prenatal care may occur for several distinct reasons, including lack of availability, financial and other barriers to utilization, attitudes and beliefs about the need and value of care, and recognition and acceptance of pregnancy (6, 9–12). It is possible that women receiving no prenatal care are not a homogenous group but rather represent several discrete subgroups, each with distinctive characteristics reflecting specific reasons behind lack of prenatal care use and associated with divergent birth outcomes.

The purpose of this investigation was to: 1) differentiate subgroups or clusters of women with similar socio-demographic and medical risk characteristics among the larger group of women receiving no prenatal care in the United States; 2) assess and validate the practical usefulness of cluster analysis to designate membership in the no-care subgroup clusters to produce a risk profile; and 3) compare birth outcomes (indicators of birth weight, gestational age, fetal growth, and infant mortality) among the subgroup clusters and with those of US women who received any prenatal care. Traditionally used in marketing to segment consumer subgroups and to guide the development and targeting of successful niche campaigns, cluster analysis may prove useful for identifying subgroups within large public health data sets. Such findings, by increasing our understanding of the unique characteristics associated with women comprising prenatal "no care" subgroups, may lead to the development of more sensitive interventions targeting their needs.

METHODS

Sample

We used records of White, Black, and Hispanic ($n = 126, 220$) US resident women delivering a live, singleton infant, and reporting no prenatal care from the 1995–1997 National Center for Health Statistics' US Linked Live Birth-Infant Death files (13). No care was determined as either having a "zero" entered for the number of prenatal care visits or having a "zero" or "blank" entered for the month care began.

Measures and Statistical Analyses

Cluster analysis, an exploratory data analysis tool that sorts cases into groups, was used to define clusters, or subgroups, of women with similar socio-demographic, behavioral, and medical risk characteristics. The following maternal risk characteristics, available on vital records, were used for the cluster analysis: age, race, marital status, education, parity, nativity/maternal birthplace, urban/suburban/rural residence, tobacco use, alcohol use, hypertension, and diabetes. Categorical variables were dummy coded as interval-level variables for use in the SAS FASTCLUS procedure (14).

PROC FASTCLUS uses a method called *nearest centroid sorting* in which a set of points called *cluster seeds* is selected as a first guess of the means of the clusters. Each observation is assigned to the nearest seed to form temporary clusters. The seeds are then replaced by the means of the temporary clusters, and the process is repeated until no further changes occur in the clusters. (15). The number of clusters or groups to be formed, termed *cluster solutions* (e.g. a five cluster solution), is specified by the analyst. Four to eight cluster solutions were examined; and a six-cluster solution was selected for producing the greatest number of meaningful clusters with at least 1000 women in each subgroup. Discriminant analysis was used to assess practical usefulness of the clustering strategy by attempting to replicate cluster assignment and determine the percentage of correct classification.

Birth outcomes examined included: very-low birth weight (<1500 g), low-birth weight (<2500 g), very preterm birth (<33 weeks gestation), preterm birth (<37 weeks gestation), small-for-gestation age (SGA: 10th percentile of birth weight for gestational age based on US reference) (16), term-SGA (10th percentile of birth weight for gestational age of infants 37–41 weeks gestation), neonatal mortality (<28 days), postneonatal mortality (28 days to <1 year), and infant mortality (<1 year). Logistic regression was used to calculate odds ratios and confidence intervals to compare risks for the selected birth outcomes among no-care subgroup clusters and the total population that received any prenatal care. The largest subgroup cluster was chosen *ad hoc* as the reference group for the logistic regression analysis.

RESULTS

Table I displays maternal socio-demographic and medical risk profiling characteristics by each of the six no-care cluster subgroups, the total no-care group ($n = 126{,}220$) and the total population of White, Black, and Hispanic women selected for this

Table I. Cluster Analysis: Maternal Profiling Characteristics, Women Receiving No Prenatal Care in the United States, 1995–1997

	Cluster 1[a]	Cluster 2[b]	Cluster 3[c]	Cluster 4[d]	Cluster 5[e]	Cluster 6[f]	Total no care	Total population
Number	7832	2222	9155	62, 014	22, 051	22, 946	126, 220	10, 610, 715
(%)	(6)	(2)	(7)	(49)	(18)	(18)	(100)	(100)
(% class correct)[g]	(100)	(100)	(95.4)	(95.7)	(83.1)	(92.4)	(93.1)	
Race (%)								
White	69	7	4	35	46	24	34	65
Black	22	7	3	53	31	25	38	16
Hispanic	9	86	93	12	23	51	28	19
Foreign-born (%)	16	89	91	7	20	35	22	16
Age (percent)								
≤20 years	2	26	26	26	28	51	29	18
21 ≤34 years	75	57	65	64	62	44	61	71
35+ years	22	17	9	10	10	5	10	11
Married (%)	65	43	46	22	32	27	29	67
Education years	16.0	0.8	5.6	11.4	12.8	8.7	10.8	12.7
Residence (%)								
Urban	23	33	38	36	35	41	36	22
Suburban	44	46	41	30	34	30	32	37
Rural	32	21	21	34	31	29	32	32
Parity (%)								
High	5	19	14	19	5	20	15	3
Average	54	46	59	68	36	41	56	56
Primipara	43	35	27	14	59	39	29	41
Diabetes (%)	1.7	1.1	0.7	0.6	1.3	0.5	0.8	2.5
Hypertension (%)	4.1	3.2	2.6	3.2	5.0	3.1	3.5	4.4
Smoke (%)	10.2	4.1	3.8	32.3	13.2	17.8	22.4	11.2
Alcohol (%)	5.0	1.1	1.2	9.2	4.3	3.8	6.4	1.2

[a]Cluster 1: older, married, White, highly educated, high risks.
[b]Cluster 2: young, foreign-born Hispanic, no education, low risks.
[c]Cluster 3: very young, foreign-born Hispanic, low education, low risks.
[d]Cluster 4: young, Black, low education, very high risks.
[e]Cluster 5: young, White, primaparas, some college, high risks.
[f]Cluster 6: very young, urban Hispanic, some high school, high risks.
[g]Percentage of cases correctly classified in cluster by discriminant analysis.

study ($n = 10,610,715$). The total no-care group was characteristically different from the total population. In contrast to the total population, women receiving no care were more likely to be Black or Hispanic, unmarried, younger, less educated, foreign born, multiparous, and urban dwelling. The no care group had lower rates of diabetes and hypertension; however, percentages of mothers reporting tobacco or alcohol use were markedly greater in the no-care group compared with the total population.

The Clusters

Six distinct no-care clusters emerged from the cluster analysis (Table I). Using discriminant analysis with the socio-demographic and medical risk characteristics as the predictors and the clusters as the dependent variable to establish if cluster membership could be replicated using discriminant functions, 93% of the no-care group was correctly classified. With the exception of Cluster 5 (18% of the sample), all clusters were correctly classified at a rate greater than 92%, and 96% of the largest cluster (Cluster 4, 49% of the sample) was correctly classified.

Cluster 1: Six percent of the no-care group was comparably older (98% ≥21 years) than the other clusters, with one quarter of the cluster aged 35 years or older. Relatively, they were highly educated (16 years), suburban dwelling (44%), and largely White (69%). Members of Cluster 1 were most likely to be married (65%), and reported the highest proportion of diabetes.

Cluster 2: Two percent of the no-care group was notably foreign born Hispanic (89%) with no education (0.8 years). Three-quarters were older than age 20, almost half were married, and they were largely suburban dwelling with low medical and behavioral risks.

Cluster 3: Seven percent of the no-care group was also notably foreign born Hispanic (91%), mostly older than age 20. Almost half were married, and they had completed elementary grades (average 5.6 years). They were suburban or urban dwellers and reported low medical and behavioral risks.

Cluster 4: Forty-nine percent of the no-care group was more likely to be Black than White (53% and 35%, respectively), largely unmarried (78%), had less than high-school education (average 11.4 years), were most likely to be average parity (68%), and least likely to be primaparous (14%). This group reported low-medical risks but extremely high behavioral risks (32.3% smoked and 9.2% drank alcohol).

Cluster 5: Eighteen percent of the no-care group was older (>72% were older than 20 years), most likely to be primaparous (59%) or average parity (36%), more White than Black (46% vs. 31%), not married (68%), and reported having more than high-school education (average 12.8 years).

Cluster 6: Eighteen percent of the no-care group was more Hispanic (51%) than White or Black (24% and 25%, respectively), very young (51% age 20 or younger) and not married (73%), with little education (average 8.7 years), urban dwelling (41%), and average parity (41%) or primaparous (39%); however, Cluster 6 had the highest rate of high parity for age (20%). Diabetes (0.5%) was rarely noted among women in Cluster 6; however, they reported the second highest rate for smoking (17.8%).

Birth Outcomes

Table II displays birth outcomes for the six no-care clusters, the total no care group, and the total population of single live births to White, Black, and Hispanic women. Overall, birth outcomes for the no-care group were two to four times worse for every measure compared to outcomes for the total population. For most birth outcomes, there was an approximate two-fold difference in risk between the cluster with the best outcomes (typically Cluster 3) and the worst outcomes (Cluster 4). The total population had more advantageous birth outcomes, compared to the no-care subgroup. For example, the rate for low-birth weight (LBW) in the total population was 6.07% compared with 20.84% for the total no-care group. Among no-care clusters, LBW varied from a low of 10.97% for Cluster 3 (very young, uneducated, foreign born Hispanics with relatively low medical

Table II. Birth Outcomes by No-Care Clusters, The Total No-Care Group, and the Total Population

	Cluster and percentage							
Birth outcome	Cluster 1[a] 7,832 (6%)	Cluster 2[b] 2,222 (2%)	Cluster 3[c] 9,155 (7%)	Cluster 4[d] 62,014 (49%)	Cluster5[e] 22,051 (18%)	Cluster 6[f] 22,946 (18%)	Total no-care 126,220 (1.2%)	Total population 10,610,715 (100%)
VLBW (%)	5.07	4.58	2.71	7.12	6.57	5.35	6.21	1.12
LBW (%)	14.80	13.83	10.97	24.59	19.34	18.78	20.84	6.07
VPT (%)	7.18	6.59	4.87	11.54	9.62	8.93	9.89	1.82
PT (%)	19.65	22.01	17.60	30.60	25.12	25.36	26.91	9.63
SGA (%)	12.77	16.63	13.42	20.78	16.94	17.82	18.46	9.43
Term SGA (%)	9.13	12.60	11.00	15.43	12.45	18.23	13.87	8.13
Infant mortality rate[g]	2.89	3.20	1.56	3.60	3.53	2.73	3.23	0.65
Neonatal mortality rate[g]	2.48	2.52	1.10	2.54	2.93	2.03	2.40	0.41
Post-neonatal mortality rate[g]	0.41	0.68	0.46	1.07	0.60	0.70	0.83	0.24

Note. VLBW = very low birth weight; LBW = low birth weight; VPT = very preterm birth; PT = preterm birth; SGA = small for gestational age; and Term SGA = term small for gestational age.
[a]Cluster 1: older, married, White, highly educated, high risks.
[b]Cluster 2: young, foreign-born Hispanic, no education, low risks.
[c]Cluster 3: very young, foreign-born Hispanic, low education, low risks.
[d]Cluster 4: young, Black, low education, very high risks.
[e]Cluster 5: young, White, primaparas, some college, high risks.
[f] Cluster 6: very young, urban Hispanic, some high school, high risks.
[g]Rates are based on number of deaths per 1000 live births.

and behavioral risks) to a high of 24.59% for Cluster 4 (young, urban, Black, low education, relatively high risks). Preterm birth occurred at a rate of 9.63% for the total population compared with 26.91% for the no-care group and ranged among the clusters from a low of 17.6% for Cluster 3 to 30.6% for Cluster 4. Small for gestational age (SGA) births were twice as frequent among women receiving no-care than for the total population, and the rate of infant mortality was almost five times greater.

Using Cluster 4, the largest cluster, as the reference group, Table III presents odds ratios and 95% confidence intervals derived from logistic regression analyses. Between Cluster 3, which generally had the best outcomes, and Cluster 4, the reference group that also generally had the worst outcomes, there were significant differences in risk for every birth weight, gestational age, fetal growth, and mortality indicator. Clusters 1–3, 5 and 6 all had a significantly lower risk of LBW than Cluster 4 (as indicated by confidence intervals that did not include 1), and the LBW odds ratio confidence intervals of Clusters 3, 5, and 6 did not overlap with any other cluster.

Table IV provides results of the logistic regression analysis of birth outcomes that include the six no-care clusters and women with any prenatal care, distinguished by race/ethnicity. White, any prenatal care women were used as the reference group. For neonatal mortality and birth weight, gestational age, and fetal growth indicators, Hispanic, any care women had slightly higher risks than White, any care women and Black, any care women had even higher risks. Although generally the no-care clusters had the highest risks, odds ratios of no-care Cluster 3 were similar to the Black any care group for LBW and very-low birth weight (VLBW) and actually lower for SGA and term-SGA. Cluster 4 had the highest odds ratios for nearly every adverse birth outcome.

DISCUSSION

Results of this study not only support earlier findings that women who receive no prenatal care are characteristically different from women who receive any care (1, 3, 6, 10), but add new insights to

Table III. Logistic Regression Odds Ratios and 95% Confidence Intervals for Birth Outcomes by Cluster, Women With No Prenatal Care in the United States, 1995–1997

	Cluster and percentage					
Birth outcome	Cluster 4[d] 62,014 (49%)	Cluster 1[a] 7,832 (6%)	Cluster 2[b] 2,222 (2%)	Cluster 3[c] 9,155 (7%)	Cluster 5[e] 22,051 (18%)	Cluster 6[f] 22,946 (18%)
LBW	1.0	0.53	0.49	0.38	0.74	0.71
	Reference	(0.50–0.57)	(0.44–0.56)	(0.35–0.40)	(0.71–0.76)	(0.68–0.74)
VLBW	1.0	0.70	0.63	0.36	0.92	0.74
	Reference	(0.63–0.77)	(0.52–0.77)	(0.32–0.42)	(0.86–0.98)	(0.69–0.79)
PT	1.0	0.56	0.64	0.48	0.76	0.77
	Reference	(0.52–0.59)	(0.57–0.71)	(0.46–0.51)	(0.73–0.80)	(0.74–0.80)
VPT	1.0	0.59	0.54	0.39	0.82	0.75
	Reference	(0.53–0.65)	(0.45–0.65)	(0.35–0.43)	(0.77–0.86)	(0.71–0.79)
SGA	1.0	0.56	0.76	0.59	0.78	0.83
	Reference	(0.52–0.60)	(0.68–0.86)	(0.56–0.63)	(0.75–0.81)	(0.79–0.86)
Term	1.0	0.56	0.79	0.68	0.78	0.89
SGA	Reference	(0.51–0.60)	(0.70–0.90)	(0.63–0.73)	(0.75–0.82)	(0.85–0.93)
Infant	1.0	0.80	0.88	0.43	0.98	0.75
Mortality	Reference	(0.69–0.91)	(0.70–1.12)	(0.36–0.50)	(0.90–1.06)	(0.69–0.82)
Neonatal	1.0	0.98	0.99	0.43	1.16	0.80
Mortality	Reference	(0.84–1.14)	(0.76–1.30)	(0.35–0.53)	(1.06–1.27)	(0.72–0.88)
Post Neonatal	1.0	0.38	0.63	0.43	0.56	0.66
Mortality	Reference	(0.27–0.54)	(0.38–1.05)	(0.31–0.59)	(0.46–0.67)	(0.55–0.78)

Note. LBW = low birth weight; VLBW = very low birth weight; PT = preterm birth; VPT = very preterm birth; SGA = small for gestational age; and Term SGA = term small for gestational age.
[a]Cluster 1: older, married, White, highly educated, high risks.
[b]Cluster 2: young, foreign-born Hispanic, no education, low risks.
[c]Cluster 3: very young, foreign-born Hispanic, low education, low risks.
[d]Cluster 4: young, Black, low education, very high risks.
[e]Cluster 5: young, White, primaparas, some college, high risks.
[f] Cluster 6 = very young, urban Hispanic, some high school, high risks.

Table IV. Logistic Regression Odds Ratios and 95% Confidence Intervals for Birth Outcomes by No-Care Clusters and any Care Racial/Ethnic Groups in the United States, 1995–1997

Birth outcome	Hispanic Any care	Black Any care	Cluster 4[d]	Cluster 1[a]	Cluster 2[b]	Cluster 3[c]	Cluster 5[e]	Cluster 6[f]
LBW	1.09	2.45	6.42	3.42	3.16	2.42	4.78	4.55
	(1.08–1.10)	(2.11–2.47)	(6.30–6.54)	(3.21–3.64)	(2.80–3.57)	(2.27–2.59)	(4.56–4.88)	(4.40–4.71)
VLBW	1.18	3.19	9.87	6.87	6.18	3.59	9.06	7.27
	(1.16–1.20)	(3.15–3.24)	(9.56–10.19)	(6.21–7.61)	(5.06–7.55)	(3.17–4.08)	(8.59–9.56)	(6.86–7.71)
PT	1.25	2.10	5.12	2.84	3.27	2.48	3.90	3.94
	(1.24–1.25)	(2.09–2.11)	(5.02–5.21)	(2.68–3.01)	(2.94–3.64)	(2.34–2.62)	(3.78–4.02)	(3.82–4.07)
VPT	1.26	3.12	10.19	6.04	5.51	4.00	8.32	7.66
	(1.24–1.27)	(3.09–3.16)	(9.93–10.47)	(5.52–6.61)	(4.61–6.59)	(3.63–4.42)	(7.94–8.72)	(7.30–8.03)
SGA	1.19	2.11	3.06	1.71	2.33	1.81	2.38	2.53
	(1.18–1.20)	(2.10–2.12)	(3.00–3.12)	(1.60–1.83)	(2.07–2.62)	(1.70–1.92)	(2.30–2.47)	(2.44–2.62)
Term	1.19	2.01	2.47	1.36	1.95	1.67	1.92	2.18
SGA	(1.18–1.19)	(2.00–2..02)	(2.41–2.52)	(1.26–1.47)	(1.72–2.21)	(1.57–1.78)	(1.85–2.00)	(2.10–2.27)
Infant	1.02	2.29	7.28	5.79	6.43	3.09	7.12	5.46
Mortality	(0.99–1.04)	(2.25–2.33)	(6.97–7.60)	(5.07–6.61)	(5.07–8.14)	(2.62–3.65)	(6.62–7.65)	(5.04–5.92)
Neonatal	1.05	2.39	8.31	8.11	8.26	3.56	9.64	6.61
Mortality	(1.02–1.08)	(2.33–2.44)	(7.90–8.76)	(7.03–9.36)	(6.33–10.77)	(2.93–4.34)	(8.90–10.44)	(6.02–7.25)
Post neonatal	0.96	2.12	5.41	2.06	3.41	2.31	3.02	3.55
Mortality	(0.93–1.00)	(2.06–2.18)	(5.00–5.88)	(1.45–2.91)	(2.05–5.67)	(1.71–3.13)	(2.54–3.59)	(3.03–4.14)

Note. The Reference Group is White, Any Care. LBW = low birth weight; VLBW = very low birth weight; PT = preterm birth; VPT = very preterm birth; SGA = small for gestational age; and Term SGA = term small for gestational age.
[a]Cluster 1: older, married, White, highly educated, high risks.
[b]Cluster 2: young, foreign-born Hispanic, no education, low risks.
[c]Cluster 3: very young foreign-born Hispanic, low education, low risks.
[d]Cluster 4: young, Black, low education, very high risks.
[e]Cluster 5: young, White, primaparas, some college, high risks.
[f]Cluster 6: very young, urban Hispanic, some high school, high risks.

this aspect of care, because we determined that the no-care women in this study were not a homogenous group. Using cluster analysis, six distinct no care clusters, or subgroups, of women with similar socio-demographic, medical, and behavioral characteristics were differentiated. Birth outcomes varied markedly among the clusters and were significantly (two to four times) worse for no-care women compared to any care women. The worst outcomes were noted for young Black women with low education and a very high proportion of maternal risk characteristics (Cluster 4). Best outcomes were noted for young, foreign-born Hispanics with lower risk factor levels (Cluster 3).

Women who receive no prenatal care in the US previously have been characterized as more likely to be older, Black, foreign-born, unmarried, to have high parity, low education, and to live in urban areas (1). Overall, no-care women in this study generally had increased parity, low education, and greater likelihood of foreign-birth (1, 3, 6, 10). However, no-care women in this study were younger and only slightly more likely to be Black than White (38% compared to 34%). Whites and Hispanics together (63% of the no-care sample) outnumbered Blacks. Whereas 22% of the no-care group was foreign-born, births to foreign-born no-care women were concentrated in two, highly defined Hispanic clusters, Clusters 2 and 3; women in all other clusters were largely born in the United States. Changing immigration patterns may explain some differences noted in our results as well as those of earlier studies. Increases in proportions of births to foreign-born women, particularly Hispanics, with a concurrent slowing of births to Whites and Blacks have been reported in recent years (17). Currently one-fifth of all US births are to foreign-born women, and one-third of all births are to unmarried women (17, 18). Finally, although fewer US teens are becoming pregnant, teens who do so are choosing to terminate their pregnancy less frequently and for various reasons some do not access prenatal care services (19).

Identification of the older, highly educated, White, married, suburban and largely primiparous Cluster 1 was unexpected. We surmise this subgroup was obscured within the larger White group in previous studies. As traditionally recognized barriers to care would not logically apply to these seemingly more advantaged women, they may have denied or failed to recognize signs of the pregnancy, or perhaps they were uninsured.

In this study, White and Hispanic clusters displayed better birth outcomes than the largely Black, Cluster 4. Consistent with other reports indicating that foreign-born mothers have better birth outcomes than their US born racial/ethnic counterparts, (17, 20, 21), more advantageous birth outcomes were consistently noted for Clusters 2 and 3, the foreign-born Hispanic clusters. Earlier studies confirm that Hispanic birth outcomes are comparable with, and in some cases better than, those of the traditional White reference group. This phenomenon, labeled by some as the "Hispanic paradox," suggests that despite low income and education, Hispanic women enjoy some comparative health advantages over other minorities, although recent works have challenged this notion, asserting that traditional epidemiologic and demographic approaches produced selection biases in earlier studies, thereby perpetuating the belief that a paradox exists when it may be due to gestational age recording errors rather than existence of innate advantages associated with race or ethnicity (22–24).

The finding of this study of most concern is the two to four-fold increased risk for negative birth outcomes demonstrated for every no-care cluster compared to women receiving any care and most particularly the excessive risks, e.g., six-fold greater infant mortality risk noted for infants of women in Cluster 4 (young, largely Black, low education, high-risk behaviors). Although White and Hispanic clusters displayed better outcomes than women in Cluster 4, their risk profiles were also notably better. The risk profile that emerged for Cluster 4 reveals a subset of women with several traditional risk factors long noted to contribute to negative birth outcomes (i.e., unmarried, low education, and extremely high behavioral risks). In contrast to Cluster 4, the older, White, highly educated Cluster 1 also reported high medical and behavioral risk characteristics, yet their outcomes were not substantially different than those of the clusters comprised of younger women with lower risk profiles. It is assumed that Cluster 1 women were healthier and generally more advantaged before becoming pregnant. Ultimately, the excessive risk for poor birth outcomes observed for Cluster 4 is likely related to their social and economic circumstance and to their risky health behaviors.

No-care women with the worst outcomes report extremely high behavioral risks even though effective interventions exist to address risks such as cigarette smoking and alcohol consumption. As one strategy for increasing visibility of women at risk for no-care, primary-care providers should screen young women for risky behaviors and appropriate counseling and interventions should be offered during the preconception and prenatal periods. Women who are foreign-born continue to be at risk for obtaining no-care, and these analyses strongly suggest that policies, such as the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (Pub. L. 104–193), (25), which limits care to some women, should be revised.

Using cluster methodology to create meaningful subsets from large public health datasets demonstrated unique benefits. To employ logistic regression with predictors similar to those produced by cluster analysis in this study would require exploring multiple models, perhaps using different reference groups and involving sophisticated interaction terms. Using the more traditional risk factor approach, the identification of the older, White, highly educated cluster obscured within the larger reference group is not likely. Widely used in business and marketing venues to produce meaningful customer behavior profiles, which are then used to guide targeted messages and to refine service delivery strategies (26–28), cluster analysis may similarly offer clues to guide development of more efficient and efficacious public health messages and outreach services to target women at risk for no-care. For example, a tourism group developed a highly successful strategy for attracting tourists by marketing "Made to Order" vacation packages targeting five, characteristic consumer clusters (27). Such an approach could be used to guide development of targeted public health messages for older, highly educated, suburban White women or for younger, less educated, urban Black or Hispanic women. As such, the identification of distinct and meaningful clusters, which were then replicated using discriminant analysis, is an advantage of the approach.

The study has several limitations, including being restricted to maternal risk characteristics available on vital records. Shortcomings associated with self-reporting and coding of secondary, vital records data, particularly birth certificate data, have been documented (29–32). Dichotomous coding of profiling characteristics as either "present" or "else" in this study possibly produced an underestimation of maternal characteristics and risks. Moreover, other variables of interest, including pregnancy intendedness and health insurance status, were unavailable. The study population was restricted to White, Black and Hispanic women. We did not adjust for multiple significance testing and, given the large sample size, most non-zero differences would probably be statistically significant whether or not they were of practical or clinical significance. Nevertheless, given the magnitude of many of the effects (as indicated by two to four-fold differences), we feel these results have practical significance.

SAS FASTCLUS is designed specifically for use with large datasets; therefore, utility for analysis of regional or state-level data is not known (14). Establishing the optimal number of clusters required multiple runs to produce meaningful groups and, although cluster membership may predict differences in birth outcomes, little of the total variation in birth outcomes in the population can be explained by clustering group. Further, some limitations to generalizability associated with grouping strategies such as cluster analysis exist. Conclusions about individuals cannot be drawn from analyses of group data (i.e., ecological fallacy), and results of this study may be less relevant to clinicians. Although discriminant analysis results indicated a high degree of group assignment accuracy, cluster analysis does not produce purely defined groups and 7% of the population was misclassified.

Cluster analysis was effective in identifying distinctive no-care subgroups that differed appreciably in terms of risks for adverse birth outcomes. Compared with more traditional methods of identifying risks associated with individual variables, this analytic approach offers an alternative and effective means of describing subgroup characteristics that can be used for public health education, intervention, and outreach efforts. Cluster profiles may help identify distinct subgroups of women most at risk for no-care and adverse birth outcomes—women who were previously invisible to the system because they had no point of entry. This is particularly relevant in an era of economic downturn and healthcare budget constraints, when there continues to be a pressing need to assure that all women receive prenatal care. No-care cluster profiles can be used to develop and focus test culturally sensitive, language and literacy appropriate messages targeting specific high risk subgroups (e.g., young Black women with low education). Finally, as different regional no-care clusters may exist and as US demographics continue to shift toward greater racial/ethnic diversity, this approach needs to be explored using state-level data and other minority groups.

REFERENCES

1. Elam-Evans L, Adams M, Gargiullo P, Kiely J, Marks J. Trends in the percentage of women who received no prenatal care in the United States, 1980–1992: Contributions of the demographic and risk effects. *Obstetr Gynecol* 1996;87:575–80.
2. Kessner DM, Singer J, Kalk CW, Schlesinger ER. *Infant Death: An Analysis by Maternal Risk and Health Care.* In: *Contrasts in health status, Vol I.* Washington DC: Institute of Medicine, National Academy of Sciences, 1973.
3. Elam-Evans L, Adams M, Gargiullo P, Kiely J. Heterogeneity between women who received prenatal care in the third trimester and those who received no prenatal care. *J Am Med Womens Assoc* 1995;50:175–7.
4. Alexander G, Kotelchuck M. Quantifying the adequacy of prenatal care: A comparison of indices. *Public Health Rep* 1996;111:408–18.
5. Alexander G, Cornely D. Prenatal care utilization: Its measurement and relationship to pregnancy outcome. *Am J Prevent Med* 1987;3:243–53.
6. Petersen DJ, Alexander GR, D'Ascoli P, Oswald J. Prenatal care utilization in Minnesota: Patterns of concern and areas for improvement. *Minn Med* 1994;77:41–5.
7. Showstack J, Budett, P, Minkler D. Factors associated with birth weight: An exploration of the roles of prenatal care and length of gestation. *Am J Public Health* 1984;74:1003–8.
8. Kogan MD, Alexander GR, Mor JM, Kieffer EC. Ethnic-specific predictors of prenatal care utilization in Hawaii. *Paediatr Perinat Epidemiol* 1998;12(2):152–62.
9. Alexander G, Korenbrot C. The role of prenatal care in preventing low birth weight. *Future Children*. 1995;5:103–20.
10. Institute of Medicine Committee to Study Outreach of Prenatal Care, Brown S, Editor. *Prenatal Care: Reaching Mothers, Reaching Infants.* Washington, DC: Institute of Medicine, National Academy Press, 1988.
11. Joyce K, Diffenbacher G, Greene J, Sorokin Y. Internal and external barriers to obtaining prenatal care. *Soc. Work Health care* 1983;9:89–96.
12. Cooney JP. What determines the start of prenatal care? Prenatal care, insurance and education. *Med Care* 1985;23:986–97.
13. National Center for Health Statistics. *1995, 1996, 1997 Birth Cohort Linked Birth/Infant Death Data Sets, NCHS CD-ROM Series 20*, Numbers 12a, 14a, 15a, ASCII Version, US Dept Health and Human Services, Centers for Disease Control and Prevention. Available at www.cdc.gov/nchs.
14. *SAS Users Guide*. Version 6, Vol. 1. SAS Institute, Cary, NC: 1989.
15. *SAS OnlineDoc, Version 8*. Cary, NC: SAS Institute, 1999. Retrieved on August 6, 2004 from http://v8doc.sas.com/sashtml/.
16. Alexander GR, Himes JH, Kaufman R, Mor JM, Kogan MD. A US national reference for fetal growth. *Obstetr Gynecol* 1996;87(2):163–8.
17. State-specific trends in US live births to women born outside the 50 states and the District of Columbia—United States, 1990 and 2000. *MMWR Morbid Mortal Weekly Rep* 2002;51:1091–5.
18. MacDorman M, Minino A, Strobino D, Guyer B. Annual summary of vital statistics–2001. *Pediatrics* 2002;110:1037–52.

19. Alan Guttmacher Institute. Teen sex and pregnancy. 1999. Retrieved July 17, 2002, from http://www.agi_usa.org/pubs/fb_teen_sex.html.
20. Rosenberg K, Desai R, Kan J. Why do foreign-born blacks have lower infant mortality than native-born blacks? New directions in African American infant mortality research. *J Natl Med Assoc* 2002;94:770–8.
21. Cervantes A, Keith L, Wyshak G. Adverse birth outcomes among native-born and immigrant women: replicating national evidence regarding Mexicans at the local level. *Matern Child Health J* 1999;3(2):99–109.
22. Palloni A, Morenoff J. Interpreting the paradoxical in the Hispanic paradox: demographic and epidemiologic approaches. *Ann. NY Acad Sci* 2001;954:140–74.
23. Buekens P, Notzon F, Kotelchuck M, Wilcox A. Why do Mexican Americans give birth to few low-birth weight infants? *Am J Epidemiol* 2000;152:347–51.
24. Franzini L, Ribble J, Keddie A. Understanding the Hispanic paradox. *Ethnicity Dis* 2001;14:496–518.
25. Centers for Medicare and Medicaid Services. *Eligibility limitations of immigrants for Medicaid and the State Children's Health Insurance Program.* Retrieved on February 25, 2003 from http://cms.hhs.gov/about/.
26. Clustan. Clustering applications. Retrieved on July 17, 2002, from http://www.clustan.com/clustering_applications.html.
27. Hudson S, Ritchie B. Understanding the domestic market using cluster analysis: A case study of the marketing efforts of Travel Alberta. *J Vacation Market* 2002;8:263–76.
28. Mochis G, Bellenger D, Curasi C. Financial service preferences and patronage motives of older consumers. *J Financ Serv Market* 2003;7:331–40.
29. Piper J, Mitchell E, Snowden M, Hall C, Adams M, Taylor P. Validation of the 1989 Tennessee birth certificates using maternal and newborn hospital records. *Am J Epidemiol* 1993;137:758–68.
30. Parrish K, Hold V, Connell F, Williams B, LoGerfo P. Variations in the accuracy of obstetric procedures and diagnoses on birth records in Washington State, 1989. *Am J Epidemiol* 1993;138:119–27.
31. Watkins M, Edmonds L, McClearn A, Mullins L, Mulinare J, Khoury M. The surveillance of birth defects: The usefulness of the revised US standard birth certificate. *Am J Public Health* 1996; 86:731–4.
32. Alexander G, Petersen D. Threats to accurate interpretation of secondary data. In: Kotch J (editor), *Maternal and child health: Programs, problems, and policy in public health.* Gaithersburg, MD, Aspen, 1997;395–403.