Maternal and Child Health Journal, Vol. 7, No. 2, June 2003 (© 2003)

A Risk Assessment Screening Test for Very Low Birth Weight

Ralitza V. Gueorguieva, PhD,[1,7] Neena P. Sarkar, PhD,[2] Randy L. Carter, PhD,[3] Mario Ariet, PhD,[4] Jeffrey Roth, PhD,[5] and Michael B. Resnick, EdD[5,6]

Objectives: To develop a risk-assessment screening tool for very low birth weight (VLBW) and to compare our empirically derived tool to the nonempirically derived screening tool used by the State of Florida. *Methods*: Birth records from the State of Florida Vital Statistics between 04/01/92 and 12/07/94 were matched with State Healthy Start prenatal records, reported from 04/01/92 through 03/31/94. Known and additional potentially important risk factors were identified from both sources. Generalized Linear Modeling techniques were used to estimate associations between risk factors and VLBW. A risk assessment system was then developed using the estimated model. The resulting screening test was compared with the one used by the Florida State Department of Health in terms of sensitivity and specificity on an independent validation sample. *Results*: The proposed screening tool had comparable specificity to the Healthy Start screening tool but significantly better sensitivity. Both instruments are simple and easy to implement. *Conclusions*: Identification of women at high risk for VLBW would be improved using the model-based screening tool developed in this paper. Public health policy makers should use statistical methods in addition to expert opinion to improve existing risk assessment methods. The actual value of an improved screening instrument is dependent on the availability of effective intervention programs.

KEY WORDS: pregnancy outcome; risk factors; VLBW; risk assessment; screening.

INTRODUCTION

Very low birth weight (VLBW, <1500 g) is one of the most serious of poor birth outcomes. Mortality and morbidity rates for such infants are high (1–3) and increase dramatically as birth weight declines below 1000 g (1, 4–6). Numerous studies have analyzed specific risk factors related to low birth weight (LBW, <2500 g) (7–9), preterm delivery (7, 10–13) (gestational age less than 36 or 37 weeks), perinatal mortality (7), and pregnancy complications (7, 14), some have incorporated risk factors into a risk assessment system for LBW (7, 8, 15–17) but none, to our knowledge, have developed a risk assessment system for VLBW. Most of LBW outcomes are 2000–2500 g births, which, although not positive outcomes per se, are no longer considered life threatening. Thus, accurate identification of women at risk to deliver a very low birth weight baby is of much greater interest than prediction of pregnancies resulting in low birth weight babies.

The goals of the present study are twofold: to develop and validate a risk assessment system

[1]Department of Epidemiology and Public Health, Yale University, New Haven, Connecticut.
[2]Eli Lilly and Company, Indianapolis, Indiana.
[3]Department of Statistics, University of Florida, Gainesville, Florida.
[4]Department of Medicine, University of Florida, Gainesville, Florida.
[5]Department of Pediatrics, University of Florida, Gainesville, Florida.
[6]Address reprint requests to Dr M. B. Resnick, Division of Neonatology, Department of Pediatrics, Box 100296, JHMHC, University of Florida, Gainesville, Florida 32610.
[7]Correspondence should be addressed to Dr R. Gueorguieva, 60 College Street, PO Box 208034, New Haven, Connecticut 06520-8034; e-mail: ralitza.gueorguieva@yale.edu.

for VLBW and to compare it to Florida's Healthy prenatal screening tool. Florida's Healthy Start program was initiated in April of 1992 independently from the Federal Healthy Start program and includes a legislative mandate that all pregnant women be screened for risk of adverse outcomes. Women who are found to be at high risk are given services such as smoking cessation, childbirth education, nutritional counseling, nursing assessment/counseling, psychological counseling, and care coordination. The Florida Healthy Start screening tool consists of an index of risk factors and associated weights, determined by a panel of State health professionals based mainly on expert opinion of the influence of each risk factor. Using these weights, a risk score is calculated for each pregnant woman and a cut point on the risk score scale classifies the woman in high or low risk category. This strategy is not based on an empirically derived scoring system and does not take into account interactions among risk factors.

In contrast, we focused on developing a model-based screening test for VLBW. Multivariable statistical methods such as logistic regression with a large number of potential risk factors have been used by a number of authors (9, 10, 13) to develop screening tests for LBW or preterm delivery by statistically estimating weights. But the selection of risk factors in such systems has often been based on univariable methods (10, 13) and interaction effects have usually not been considered (9, 10, 13) because of limited sample sizes. Michielutte *et al.* (16) included interactions, and tested the accuracy of screening tests for preterm LBW, term LBW, and all LBW. They concluded that, for predictive use, the distinction between preterm and term LBW was unnecessary and called for further research to identify additional modifiable risk factors to be included in improved screening tests, specifically suggesting testing variables related to life stress.

In our study we simultaneously consider several such variables in addition to other previously identified risk factors for VLBW and investigate all two-way interactions between risk factors. After careful model selection, a risk-scoring sheet similar to the one used in the Healthy Start Program is developed. The accuracy of our risk index is then tested and compared with that of Florida's Healthy Start screening tool in an independent validation sample. A discussion of the advantages and disadvantages of the two approaches concludes the paper.

METHODS

Sample Description

Data were collected by the State Health Office of Florida and consisted of Healthy Start prenatal screening records of pregnant women, and of children's birth certificate records. To obtain the sample used for development of the screening tool, we merged the 254,020 Healthy Start prenatal screening records between 04/01/92 and 3/31/94 with all of Florida's 512,891 birth certificate records, reported from 04/01/92 and 12/07/94, using mother's social security number for linkage. We matched 166,372 records (65% of the available Healthy Start screening records). Multiple births and records with missing values were excluded. To obtain the validation sample used to check the accuracy of the model based screening test, we merged the subsequent 64,640 Healthy Start prenatal screening records from 04/01/94 to 12/12/94 to the 189,636 birth certificate records from 04/01/94 to 12/31/95 using the same strategy. The final validation sample consisted of 37,250 observations (58% of the available Healthy Start screening records). Details of the matching algorithms and matching success are available from the authors. Matching rate is not critical for the methodological analyses presented in this manuscript.

We identified all potential risk factors for VLBW in the Birth Vital Statistic and the Healthy Start data sets and excluded from further consideration risk factors with large proportion of missing values. Fourteen risk factors had complete enough data to be considered for inclusion in the statistical modeling process. Six variables came from Healthy Start prenatal records: illness requiring continuous medical care, safe place to live and enough food, moved more than three times in the last 12 months, transportation problems, sexually transmitted diseases, and unwanted pregnancy; eight variables were obtained from the birth certificate records: mother's education, age, marital status, race, previous pregnancy experience, alcohol use, smoking status, and interpregnancy interval (Table I). Interpregnancy interval was defined as time between termination of the last previous pregnancy and the last menstrual period before the current pregnancy. The successfully matched records both in the model and in the validation sample were comparable to the total Healthy Start prenatal screen and birth vital statistics records on the available predictor variables (Table I, last two columns). The sample

Table I. Description of Variables, Percents of Infants Per Level of Each Variable and Rates of VLBW for Original Unmatched Data Sets (Birth Vital Statistics and Healthy Start), and for Matched Data Sets Used for Statistical Analysis (Sample Used for Model Development and Validation Sample)

		Sample used for model development ($N = 166,372$)		Validation sample ($N = 37,250$)		Birth Vital Statistics ($N = 512,891$)	Healthy Start ($N = 254,020$)
Variable	Levels	Percent of infants	Crude rate of VLBW (per 1000)	Percent of infants	Crude rate of VLBW (per 1000)	Percent of infants	Percent of infants
Mother's age	11–19	17.1	12.9	17.7	16.3	13.3	—
	35 or older	8.9	11.0	9.6	11.2	10.9	—
	20–34	74.0	8.2	72.7	9.0	75.8	—
Mother's education	<High school	27.0	11.5	27.0	13.3	23.2	—
	High school	39.5	10.0	38.3	11.0	37.8	—
	>High school	33.5	6.6	34.7	7.7	39.0	—
Marital status	Not married	40.5	12.9	41.9	14.0	34.4	—
	Married	59.5	6.8	58.1	7.5	65.6	—
Mother's race	Black	21.1	17.1	21.2	20.4	19.5	—
	Other	18.1	10.0	17.9	10.6	23.8	—
	White	60.8	6.3	60.9	7.0	56.7	—
Illness requiring continuous care	Yes	5.4	15.0	6.0	15.3	—	5.9
	No	94.6	8.9	94.0	10.2	—	94.1
Safe place to live and enough food	No	5.6	11.7	5.7	13.8	—	6.5
	Yes	94.4	9.1	94.3	10.3	—	93.5
Moved more than three times in the last year	Yes	8.7	12.5	8.6	12.7	—	9.8
	No	91.3	9.0	91.4	10.3	—	90.2
Previous pregnancy experience	No previous pregnancy	35.7	10.5	36.2	12.5	31.9	—
	Previous failure	32.6	10.9	32.2	11.8	35.4	—
	Successful pregnancy	31.7	6.3	31.6	6.9	32.7	—
Sexually transmitted diseases	Yes	3.1	13.1	3.4	11.8	—	3.4
	No	96.9	9.2	96.0 0.6	10.5	—	96.6
Transportation problem	Yes	9.6	11.3	9.7	10.0	—	11.1
	No	90.5	9.1	90.3	10.6	—	88.9
Unwanted pregnancy	Unwanted	6.7	12.7	6.4	14.2	—	7.0
	Wanted	93.3	9.0	93.6	10.2	—	93.0
Alcohol use[a] (no. of drinks per week)	0	98.9	9.2	99.0	10.5	98.9	—
	1–7	1.0	16.2	0.9	14.7	1.0	—
	8 and above	0.1	57.7	0.1	—	0.1	—
Smoking[a] (no. of cigarettes per day)	0	82.2	9.1	83.4	10.5	84.9	—
	1–12	11.3	10.5	10.7	11.5	9.5	—
	13 and above	6.5	9.9	5.9	8.6	5.7	—
Interpregnancy time in years[a]	0, no experience	35.7	10.5	36.2	12.5	34.8	—
	0.5–3.5 years	47.7	7.9	47.0	9.0	48.2	—
	4+ years	16.7	10.6	16.8	10.4	17.0	—
VLBW	(<1500 g)	0.9	—	1.0	—	1.4	—[b]

[a]Variable treated as continuous in the analysis.
[b]Percent of VLBW is not available because some of the Healthy Start prenatal screen records did not match to birth certificate records.

of records with missing data was also similar on all predictor variables to the complete matched data set. The percent of VLBW in the Birth Vital Statistics data set was higher than in our matched data (1.4% compared to 0.9% and 1.0%, last row of Table I) which can be attributed to women with no or late prenatal care (and hence no Healthy Start prenatal screen) who are at a higher risk to deliver very low birth weight infants.

Statistical Modeling

We used multivariable Generalized Linear Models (18) (PROC GENMOD in SAS/STAT

Table II. Estimates of Relative Risks of VLBW for Significant Main Effects and Significant Interactions Based on the Matched Sample Used for Model Development ($N = 166, 372$)

Factor	Level	Reference level	Relative risk	Confidence interval for relative risk
Mother's age	11–19	20–34	1.10	(0.95, 1.28)
	35 or older		1.35	(1.13, 1.62)
Mother's education	<High school	>High school	1.28	(1.10, 1.50)
	High school		1.31	(1.14, 1.49)
Marital status	Single	Married	1.42	(1.20, 1.68)
Mother's race	Black	White	2.33	(2.06, 2.65)
	Other		1.42	(1.23, 1.63)
Illness requiring continuous care	Yes	No	1.39	(1.16, 1.66)
Safe place to live and enough food	No	Yes	1.01	(0.82, 1.25)
Moved more than three times in the last year	Yes	No	1.22	(0.98, 1.51)
Previous pregnancy experience	No previous pregnancy	Successful pregnancy	1.66	(1.29, 2.13)
	Previous failure		1.69	(1.47, 1.94)
Race × Marital status	Black × Single	Black × Married	0.90	(0.71, 1.13)
	Other × Single	Other × Married	1.35	(1.05, 1.75)
	White × Single	White × Married	1.42	(1.16, 1.74)
Marital status × Moved more than three times in the last year	Single × Yes	Single × No	0.99	(0.78,1.27)
	Married × Yes	Married × No	1.54	(1.15, 2.06)
Safe place to live and enough food × Moved more than three times in the last year	No × Yes	No × No	1.91	(1.30, 2.81)
	Yes × Yes	Yes × No	1.06	(0.88, 1.27)
Alcohol use[a] (drinks per week)			1.05	(1.02, 1.08)
Smoking[b] (cigarettes per day)			1.01	(1.01, 1.02)

[a]Each additional drink per week increases the risk of VLBW by 5%.
[b]Each additional cigarette per day increases the risk of VLBW by 1%.

software) to analyze the effect of the independent variables given in Table I on the probability of VLBW. The initial model included all main effects and all estimable two-way interactions of the risk factors. Backwards elimination with a significance level of 0.04 was used to eliminate nonsignificant effects. We also considered using a significance level of 0.05 but that led to the creation of a much more complicated scoring sheet without improving predictive accuracy. Interpregnancy interval was estimated to have a change point linear relationship with the probability of VLBW. The change point (in 6 month increments) giving the best model-fit when controlling for all other significant effects was selected. The final fitted model was used to calculate adjusted relative risks (RR) of VLBW for each level of each factor relative to the lowest risk reference level.

Development of Screening Tests

To develop the Healthy Start screening test, the Healthy Start Advisory Committee selected a number of risk factors that potentially increase the risk of LBW or preterm delivery. The committee included epidemiologists, obstetricians, nurses, social workers, and policy and program representatives from key state and local maternal and child health partners. The screening criteria were selected based on literature review, professional judgment and on descriptive statistics from birth certificate data between 1989 and 1990. Each of the risk factors was divided into two categories except previous pregnancy experience that was divided into three categories. Each high-risk category then was assigned a weight based

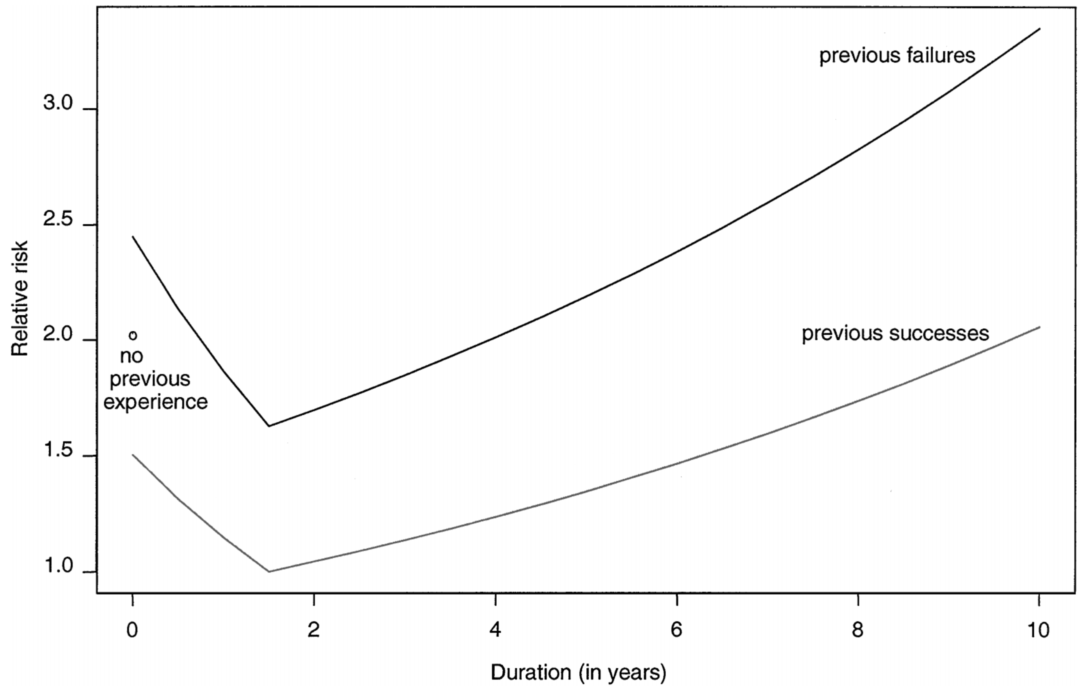

Fig. 1. Relative risk of very low birth weight by interpregnancy interval and previous pregnancy experience. (Denominator for each relative risk is the risk for women with previous successful experience and interpregnancy interval of 1.5 years.)

on expert opinion and/or unadjusted risk increase estimates. Thus a risk score could be computed for each woman by adding the weights for all high-risk categories to which that woman belonged. Finally, a cutoff value was selected so that women whose score was higher than the cutoff were considered to be at high risk and were referred for prenatal services. The cutoff was selected based on professional judgment.

In contrast, we developed a VLBW risk-scoring sheet based on the statistical model, which included interactions of risk factors and in which estimates of the effect of each predictor were adjusted for other predictors. The factors included in our scoring sheet were the significant effects in the model. The weight for a given factor level was defined as 100 times the value of the coefficient estimate for that level. The weight for a combination of factor levels was formed by adding the coefficients for the corresponding main effects and the two-way interactions, and by multiplying them by 100. The total score for VLBW could be obtained by adding the weights. Different versions of the screening test are defined by specifying different cutoffs on the total risk score.

Comparison of Healthy Start and Model-Based Screening Test

The accuracy of the model-based screening tests and of the Healthy Start screening test was measured by the percentage of all VLBW births that were predicted correctly by each test (sensitivity), the percentage of non-VLBW births that were predicted correctly (specificity) and the likelihood ratio defined as sensitivity/(1-specificity). Higher likelihood ratio corresponds to higher sensitivity and/or specificity. The caseload of any given test was the number of women who tested positive by that test. We considered several additional cutoff values for the Healthy Start screening test in order to compare the two tests over a range of caseloads and to create receiver–operating–characteristic curves (ROC curves) for the two approaches.

RESULTS

Statistical Modeling

Table II shows the relative risk estimates and associated 95% confidence intervals for the

Table III. Healthy Start Risk Scoring Sheet

Factor	Level	Weight	Contribution to risk score
Mother's age and mother's education	Younger than 18	1	+...........
	Older than 35	1	+...........
	<High school	1	+...........
	High school or more	—	0
	18–35	—	0
	<High school	1	+...........
	High school or more	—	0
Race	Black	1	+...........
	Non-black	—	0
Marital status	Not married	1	+...........
	Married	—	0
Trimester prenatal care began	Third	1	+...........
	First or second	—	0
Prepregnancy weight of 115 lbs or less	Yes	1	+...........
	No	—	0
Previous pregnancy experience	No previous pregnancy	1	+...........
	Previous failure	2	+...........
	Successful experience	—	0
Illness requiring continuous care	Yes	4	+...........
	No	—	0
Number of cigarettes per day	10 or more	1	+...........
	Less than 10	—	0
More than one drink per day	Yes	2	+...........
	No	—	0
Safe place to live and enough food	No	1	+...........
	Yes	—	0
Moved more than three times in the last year	Yes	1	+...........
	No	—	0
Transportation problems year	Yes	1	+...........
	No	—	0
Sexually transmitted disease in the last 6 months	Yes	1	+...........
	No	—	0
Unwanted pregnancy	Yes	1	+...........
	No	—	0
Total score			

factors in the final model. The effect of interpregnancy interval on the risk of VLBW is shown in Fig. 1 by the levels of previous pregnancy experience.

Development of Screening Tests

Tables III and IV describe the scoring methods of the Healthy Start screening and of the model-based approach. High scores correspond to higher risk of VLBW. The cutoff value used by Florida's Healthy Start screening test is 4. Different cutoff points can be considered on the model-based screening instrument corresponding to different caseloads.

Comparison of Healthy Start and Model-Based Screening Test

Table V shows the accuracy of the model-based and the Healthy Start screening tests, defined by various cutoff points, in the validation sample. Figure 2 shows ROC curves for the two approaches. Since the Healthy Start risk score could take only several positive integer values, there were fewer possible cutoff points than in the model-based approach. To facilitate the comparison between the two screening tools in Table V, we selected only nine cutoff points for the model-based approach that produced similar caseload to the nine cutoff points in the Healthy Start screening tool. However, the ROC curve for the model-based approach in Fig. 2 is based on five

Table IV. Model-Based Risk Scoring Sheet

Factor	Level	Weight	Contribution to risk score
Mother's age	11–19	11	+...........
	35 or older	26	+...........
	20–34	—	0
Mother's education	<High school	17	+...........
	High school	22	+...........
	>High school	—	0
Alcohol use		5 times the number of drinks per week	+...........
Smoking		Number of cigarettes per day	+...........
Illness requiring continuous care	Yes	31	+...........
	No	—	0
Previous pregnancy experience	No previous pregnancy	30	+...........
	Previous failure	49	+...........
	Successful experience	—	0
Time from last menstrual period to end of previous pregnancy	Less than 1.5 years	27 × (1.5–duration)	+...........
	1.5 years or more	9× (duration–1.5)	+...........
Marital status	Not married	42	+...........
	Moved	–2	−...........
	Not moved	—	0
	Black	72	+...........
	Other	39	+...........
	White	—	0
	Married	—	0
	Moved	45	+...........
	Not moved	—	0
	Black	119	+...........
	Other	44	+...........
	White	—	0
Safe place to live and enough food	No	–19	−...........
	Moved	53	+...........
	Not moved	—	0
	Yes	—	0
Total score			

times as many possible cutoff points as shown in Table V.

The likelihood ratio for the model-based approach was higher than the likelihood ratio for the State's approach at every caseload percentage (Table V). Specificity was similar for both methods at every caseload percentage, but sensitivity was uniformly better for the model-based screening test. Positive predictive value and negative predictive value were always higher for the model-based approach but the differences were small in absolute value. The ROC curve in Fig. 2 shows that model-based test sensitivity was up to 10% points higher than that of Healthy Start Screening tool.

DISCUSSION

The methods used in this study to develop an improved screening test were similar to those used by Michielutte *et al.* to develop a screening test for identification of women at high risk for LBW births. The accuracies of our test for identifying VLBW were similar to the accuracies of Michielutte's and deCaunes' tests for identifying LBW. Although the variables considered in these studies were quite different, the overlap of significant risk factors was notable. Because we were limited by the information available in the Birth Vital Statistics Data Set and on the Florida's Healthy Start screening form, we could not consider some important risk factors for VLBW. For example, four of the significant factors identified by Michielutte were not considered in our analysis: two because of unavailability (economic status, and uterine anomaly) and two because of frequent missing values (maternal height and weight). We also could not consider psychosocial risk factors such as anxiety and family dissatisfaction identified by Herrera (7), previous low-birth weight or preterm delivery for multiparous women

Table V. Accuracy of Florida's Healthy Start Screening Tool and of the Model-Based Screening Tool by Various Cutoff Points

Healthy start screening tool					Model-based screening tool				
Cutoff	Caseload percent	Characteristics		Likelihood ratio[a]	Cutoff	Caseload percent	Characteristics		Likelihood ratio
9	1.81	Sensitivity	3.32	1.84	237	1.89	Sensitivity	4.35	2.34
		Specificity	98.20				Specificity	98.14	
		PPV[b]	1.93				PPV	2.41	
		NPV[c]	98.97				NPV	98.98	
8	3.36	Sensitivity	5.88	1.77	222	3.25	Sensitivity	7.93	2.48
		Specificity	96.67				Specificity	96.80	
		PPV	1.84				PPV	2.56	
		NPV	98.98				NPV	99.00	
7	5.95	Sensitivity	9.97	1.69	217	5.13	Sensitivity	14.83	2.95
		Specificity	94.09				Specificity	94.97	
		PPV	1.76				PPV	3.04	
		NPV	99.00				NPV	99.06	
6	10.16	Sensitivity	16.37	1.62	207	11.15	Sensitivity	25.83	2.35
		Specificity	89.90				Specificity	89.00	
		PPV	1.69				PPV	2.43	
		NPV	99.02				NPV	99.12	
5	17.33	Sensitivity	29.16	1.69	186	17.59	Sensitivity	38.11	2.19
		Specificity	82.79				Specificity	82.62	
		PPV	1.77				PPV	2.27	
		NPV	99.10				NPV	99.21	
4	28.23	Sensitivity	45.78	1.63	155	27.76	Sensitivity	49.36	1.79
		Specificity	71.96				Specificity	72.47	
		PPV	1.70				PPV	1.87	
		NPV	99.21				NPV	99.26	
3	40.29	Sensitivity	60.10	1.50	140	41.48	Sensitivity	65.98	1.60
		Specificity	59.92				Specificity	58.78	
		PPV	1.57				PPV	1.67	
		NPV	99.30				NPV	99.39	
2	51.69	Sensitivity	67.77	1.32	117	50.38	Sensitivity	73.15	1.46
		Specificity	48.43				Specificity	49.86	
		PPV	1.38				PPV	1.52	
		NPV	99.30				NPV	99.43	
1	59.64	Sensitivity	71.87	1.21	102	58.02	Sensitivity	77.49	1.34
		Specificity	40.49				Specificity	42.18	
		PPV	1.26				PPV	1.40	
		NPV	99.27				NPV	99.44	

[a] Higher likelihood ratios indicate a better screening tool at fixed caseload.
[b] PPV is positive predictive value.
[c] NPVA is negative predictive value.

found to be highly significant by a number of authors (7, 8, 13, 14, 19–21). Such factors could be considered for inclusion in future revisions of the Healthy Start questionnaire in place of other factors that are not highly associated with the outcome, or that are colinear with other factors. Inclusion of additional factors could lead to further improvement of the model-based screening tool.

Because the choice of risk factors in the model-based screening test was based on multivariable statistical methods and because weights were derived from model parameters, one should expect the model-based screening tool to outperform the less statistically derived instrument used by the State. Indeed the selection of factors, weights and cutoff value in Florida's Healthy Start screening tool did not take into account the possible interactions of the risk factors. In contrast, our model selection approach allowed us to include only factors and interactions that had significant effects on the outcome, to select empirically justified weights and cutoff that provided good sensitivity and specificity at a manageable caseload. Statistical techniques are no substitution for thinking through the conceptual framework that exists among the predictor and response variables. Such theoretical considerations guided our selection of risk factors and interactions for inclusion in the model. The stepwise model-selection procedure helped identify the most significant effects that produced good predictions with a parsimonious model.

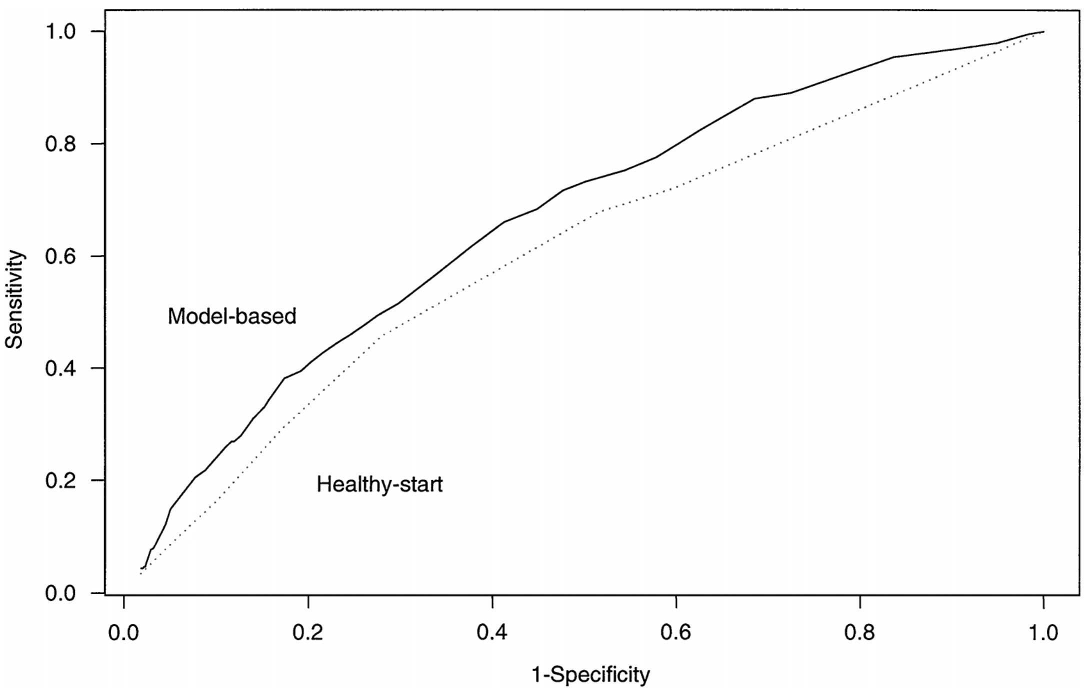

Fig. 2. Receiver–operating–characteristic curves for the model-based and for the Healthy Start screening tools. (Since VLBW is a rare occurrence, caseload percentage is approximately equal to (1-specificity) percent.)

The relative simplicity of the final model was essential as it allowed us to develop a scoring sheet nearly as simple to use as the one based on the Healthy Start screen. That is why we chose to use a significant level of 0.04 instead of 0.05 in our model-selection procedure. The scoring sheet based on a significance level of 0.05 was much more complicated. The computation of the total score in the model-based screening tool proposed herein is easy to perform with the simplest calculator. If total scores are computed by hand, computational errors can be made and the error rate of our instrument may be higher than the error rate using the Healthy Start screening. However if a computerized version of the risk assessment tool is used, total scores will be calculated by a computer, and computational problems will not occur. A pilot field study to demonstrate the usefulness of such an approach should be implemented.

We believe that the better accuracy in identifying women at risk justifies the use of a more complex weighting scheme, and that health care providers are interested in more scientific and evidence-based practices. Since new knowledge and data constantly become available, the process of derivation of the model-based screen form should be repeated periodically.

The value of being able to better predict whether a woman is at risk for adverse pregnancy outcome depends in part on the availability of interventions that are likely to improve the outcome. However interventions aimed at reducing VLBW births have not shown conclusive effectiveness (23, 24). Enhanced prenatal care, nutritional interventions and the use of tocolytics have not generally been successful, efforts to predict and prevent preterm delivery have given frustrating results (3, 13, 23), and interventions such as smoking cessation programs have been difficult to implement (22). Hence continued effort is needed to identify and refine effective practices to reduce low birth weight and to discover most appropriate methods of implementing these practices.

The advantages of the model-based approach can be demonstrated in two ways. First we fix caseload (number of women identified as being at high risk for VLBW) and compare sensitivities of the two tests. On the basis of the numbers in Table V and considering a caseload of about 28%, the sensitivity of the model-based screen is about 4% higher than the sensitivity of Florida's Healthy Start screen. Hence our test may help identify 40 additional women for every 1000 women delivering LBW babies while keeping the total number of women referred for enhanced services the same.

An alternative way to demonstrate the advantages of the model-based approach is to fix sensitivity (number of women with LBW babies correctly identified by the test) and compare the caseloads and specificities of the two tests. On the basis of the receiver operating curves in Fig. 2, a caseload of 40% is required to achieve 60% sensitivity using the Healthy Start test, while a caseload of 37% is needed to achieve the same sensitivity using the model-based test. Thus, an intervention program that uses the screening test developed in this paper would be able to identify the same number of women who will deliver VLBW babies as the Healthy Start screen but will refer for enhanced services fewer women, and hence will have lower cost.

In this time of health care reform and dwindling resources for public health programs, the need for cost effective interventions is great. Resources generally are not sufficient to provide expanded preventive prenatal services to all pregnant women although careful monitoring and response to emerging pregnancy complications for all pregnant women is advocated by many (25). Public health policy makers, therefore, should pursue opportunities to improve existing tests using the empirical methods illustrated here.

ACKNOWLEDGMENTS

This study was supported, in part, by a contract with the Florida Department of Education (University of Florida Project # 93042001). We thank Dr Charles Mahan and Daniel Thompson, MPH, for their participation in the concept and design of the study and their help in the critical revision of the manuscript. We also acknowledge Dr Leslie Beitsch, Dr Richard Bucchiarelli, Dr John Curran for their comments, and Dr Chi-Hse Teng for his help with data management and analysis.

REFERENCES

1. Resnick MB, Carter RL, Ariet M, Bucciarelli RL, Evans JH, Furlough RR, Ausbon WW, Curran JS. Effect of birth weight, race, and sex on survival of low-birth-weight infants in neonatal intensive care. *Am J Obstet Gynecol* 1989;161(1):184–7.
2. Carter RL, Resnick MB, Ariet M, Shieh G. A random coefficient growth curve analysis of development in low-birth-weight infants. *Stat Med* 1992;11:243–56.
3. Lockwood CJ. Predicting premature delivery—no easy task. *N Engl J Med* 2002;346(4):282–4.
4. Hack M, Taylor GH, Klein N, Eiben R, Schatschneider C, Mercuri-Minich N. School-age outcomes in children with birth weight under 750 g. *N Engl J Med* 1994;331:753–9.
5. Resnick MB, Stralka K, Carter RL, Ariet M, Bucciarelli RL, Furlough R, Evans JH, Curran JS, Ausbon WW. Effects of birth weight and sociodemographic variables on mental development of neonatal intensive care unit survivors. *Am J Obstet Gynecol* 1990;162:374–8.
6. Resnick MB, Roth J, Ariet M, Carter RL, Emerson JC, Hendrickson JM, Packer AB, Larsen JJ, Wolking WD, Lucas M, Schenck BJ, Fearnside B, Bucciarelli RL. Educational outcome of neonatal intensive care graduates. *Pediatrics* 1992;89:373–8.
7. deCaunes F, Alexander GR, Berchel C, Guengant JP, Papiernik E. Anamnestic pregnancy risk assessment. *Int J Gynecol Obstet* 1990;33:221–7.
8. Herrera JA, Salmeron B, Hurtado H. Prenatal biopsychosocial risk assessment and low birth weight. *Soc Sci Med* 1997;44:1107–14.
9. Ernest JM, Michielutte R, Meis PJ, Moore ML, Sharp PC: Identification of women at high risk for preterm-low-birth weight births. *Prevent Med* 1988;17:60–72.
10. Guzick DS, Daikoku NH, Kaltreider DF. Predictability of pregnancy outcomes in preterm delivery. *Obstet Gynecol* 1984;63:645–50.
11. Mercer BM, Goldenberg RL, Das A, Moawad AH, Iams JD, Meis PJ, Copper RL, Johnson F, Thom E, McNellis D, Miodovnik M, Menard MK, Caritis SN, Thurnau GR, Bottoms SF, Roberts J. The preterm prediction study: A clinical risk assessment system. *Am J Obstet Gynecol* 1996;174:1885–93.
12. Goldenberg RL, Iams JD, Mercer BM. The preterm prediction study: The value of new versus standard risk factors in predicting early and all spontaneous preterm births. *Am J Public Health* 1998; 88:233–8.
13. Mueller-Heuback E, Guzick DS. Evaluation of risk scoring in a preterm birth prevention study of indigent patients. *Am J Obstet Gynecol* 1989;160:829–35.
14. Berglund A, Lindmark G. The usefulness of initial risk assessment as a predictor of pregnancy complications and premature delivery. *Acta Obstet Gynecol Scand* 1999;78:871–6.
15. Eisner V, Brazie JV, Pratt MW, Hexter AC. Risk of low birth weight. *Am J Public Health* 1979;69(9):887–893.
16. Michielutte R, Ernest JM, Moore ML, Meis PJ, Sharp PC, Wells HB, Buesche PA. A comparison of risk assessment models for term and preterm low birth weight. *Prev Med* 1992;21:98–109.
17. Creasy RK, Gummer BA, Liggins GC. System for predicting spontaneous preterm birth. *Obstet Gynecol* 1980;55:692–5.
18. McCullagh P, Nelder JA. *Generalized linear models*. London: Chapman and Hall, 1983.
19. Malloy MH. Risk of previous very low birth weight and very preterm infants among women delivering a very low birth weight and very preterm infant. *J Perinatol* 1999;19:97–102.
20. Ekwo EE, Moawad A. The relationship of inter-pregnancy interval to the risk of preterm births to black and white women. *Int J Epidemiol* 1998;27:68–73.
21. Kogan MD, Alexander GR, Kotelchuck M, Nagey DA. Relation of the content of prenatal care to the risk of low birth weight. *JAMA* 1994;271:1340–5.
22. Goldenberg RL. The prevention of low birth weight and its sequelae. *Prev Med* 1994;23(5):627–31.
23. Iams JD, Newman RB, Thom EA, Goldenberg RL, Mueller-Heuback E, Moawad A, Sibai BN, Caritis SN, Miodovnik M, Paul RH, Dombrowski MP, McNellis D. Frequency of uterine contractions and the risk of spontaneous preterm delivery. *N Engl J Med* 2002;346(4):250–5.
24. American College of Obstetrics and Gynecology. Assessment of risk factors for preterm birth. ACOG *Practice Bulletin.* Obstet Gynecol 2001;98:709–16.
25. Moore ML, Freda MC. Reducing preterm and low birth weight births: Still a nursing challenge. *Am J Matern Child Nurs* 1998;23:200–8.