

Development and validation of an infant morbidity index using latent variable models

Xuefeng Liu^{1,*},[†],[‡] and Jeffrey Roth^{2,§}

¹*Translational Research and Clinical Epidemiology, Department of Internal Medicine at the Wayne State University, Detroit, MI 48201, U.S.A.*

²*Maternal Child Health and Education Research and Data Center, Department of Pediatrics at the University of Florida, Gainesville, FL, U.S.A.*

SUMMARY

Birth defect, abnormal condition of the newborn, developmental delay or disability and low birth weight are four major infant morbidity outcomes. Most studies have focused on assessment of the effects of risk factors on each of these outcomes or of the relationship among these outcomes or both. Little attention has been paid to the development of a composite index, which is a summary construct of infant morbidity outcomes. In this paper, we develop extended latent variable (LV) models and modified Gauss–Newton algorithms for multiple multinomial morbidity outcomes with complete responses. By assuming the marginal distribution of the LV to be log-normal, we model the conditional probability of each outcome as a nonlinear function of the LV, which has properties similar to the logistic function. The estimated generalized nonlinear least-square method is used to solve equations for parameters of interest. The models are applied to an infant morbidity data set. A new single variable, called infant morbidity index (IMI) that functions as a summary of four infant morbidity outcomes and represents propensity for infant morbidity, is developed. The validity of this index is then assessed in detail. It is shown that the IMI is correlated with each of the individual outcomes, with infant mortality and with a face-valid index of morbidity outcomes, and can be used in future research as a measure of propensity for infant morbidity. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: infant morbidity outcomes; latent variable models; modified Gauss–Newton algorithms; validation; infant morbidity index

*Correspondence to: Xuefeng Liu, Translational Research and Clinical Epidemiology, Department of Internal Medicine at the Wayne State University, 540 E. Canfield, TRACE, UHC, Detroit, MI 48201, U.S.A.

[†]E-mail: xli@med.wayne.edu

[‡]Assistant Professor.

[§]Associate Professor.

1. INTRODUCTION

Infants have varying propensities for morbidity. The propensity for morbidity (denoted by S) is manifest in morbidity outcomes. There are several different morbidity outcomes to consider, such as birth defect (BD), developmental delay or disability (DDD) and low birth weight (LBW). These outcomes are associated with each other [1–8] and with risk factors, such as poverty, smoking during pregnancy and alcohol consumption during pregnancy. Most current studies have focused on assessment of the effects of risk factors on each of these outcomes or of the relationship between different outcomes or both [1, 6–12]. It is not useful clinically to identify risk factors and quantify the strength of the relationship between risk factors and infant morbidity only by analyzing the effects of risk factors on the individual morbidity outcome. A single measurement of morbidity analogous to, say, blood pressure as a measure of cardiac output is needed. The objective of this study is to develop an index of infant morbidity (IM) combining four pregnancy outcomes, which will be described below and to assess the validity of the index.

BD, abnormal condition (AC) of the newborn, DDD and LBW are four major contributors to infant morbidity [5–14]. BD is defined as abnormal development of the fetus resulting in death, malformation, growth retardation or functional disorders. Approximately 150 000 babies are born each year in the U.S. with BDs. Approximately 3 per cent of all children born in the U.S. have a major malformation at birth. BD is one of the leading causes of infant mortality [13, 14]. The major risk factors of BD are environmental exposure, maternal alcohol consumption and maternal cigarette smoking during pregnancy [15–18]. AC of the newborn denotes infants who have anemia ($HCT < 39/HGB < 13$), birth injury, fetal alcohol syndrome, hyaline membrane disease/RDS, meconium aspiration syndrome, seizures or assisted ventilation. Developmental delay is the slowed or impaired development of a child under 5 years old. Development disability when applied to infants and young children means individuals from birth to age 5, inclusive, who have substantial developmental delay or specific congenital or acquired conditions with a high probability of a mental or physical impairment or combination of mental and physical impairments. Many children show problems of DDD with time, e.g. 6–7 per cent by 1 year of age and 12–14 per cent by school age. An outcome variable, that is related to DDD, is LBW. LBW is a strong predictor of DDD in early childhood [8, 19]. Other major causes of DDD are environmental toxins, substance abuse, poor nutrition and genetic syndromes [8–12, 19]. LBW refers to infants born less than 2500 g. Birth before 37 weeks gestation is a major component of LBW. In the United States, preterm birth accounts for 50 per cent of long-term neurologic morbidity [20].

Since these outcome variables are major causes of infant morbidity, the effects of risk factors on these outcome variables and their relationship have been analyzed in many studies [7–12, 19]. But little attention has been paid to the development of a composite index which is a summary construct of infant morbidity. Such an index, nevertheless, could provide a single comprehensive measure that might be useful in studies of infant health. Compared with the subjective face-valid index (FVI) (defined later), the new index not only identifies all the patterns of morbidity outcomes but also allows early intervention programs aimed at improving the health of infants with morbidity.

Latent variable (LV) models in latent class setting have recently been developed for a wide array of problems and applications. Garrett *et al.* [21] used the population-based latent class model as the standard of comparison for diagnoses for which there was no gold standard of diagnosis. Lin and McCulloch [22] considered a latent class model to uncover subpopulation structure for both biomarker trajectories and the probability of disease outcome in highly unbalanced longitudinal data. Qu *et al.* [23] developed a general latent class model with random effects to model the

conditional dependence among multiple diagnostic tests. In this study, we propose LV models in which the LV is not treated as categorical in the latent class setting but as continuous in the latent trait setting. By assuming the marginal distribution of the LV to be log-normal, we model the conditional distribution of each morbidity outcome as a nonlinear function of the LV. These models are utilized to combine four outcomes, BD, AC, DDD and BW, into a single infant morbidity index (IMI) that reflects the propensity for morbidity (LV). An overall assessment of the validity of the IMI is made by correlating it with each outcome, with IM and with a FVI of morbidity outcomes. In our analysis, we assume that the likelihood of each of these four outcomes is affected by an underlying propensity for morbidity that has a common influence on all four morbidity outcomes.

This paper is structured as follows. In Section 2, we describe the source and creation of infant morbidity data used in this study. Sections 3 presents extended LV models for multiple multinomial outcomes, and models for modeling conditional probabilities. In Sections 4 and 5, we introduce modified Gauss–Newton algorithms for estimating the parameters of interest in LV models, present a method for obtaining initial estimates, and test the goodness of fit of the models. We develop an IMI and a class of related validation procedures in Section 6. Several issues are discussed in the final section.

2. DATA SET

The infant morbidity data were derived from the merger of four data sources. The base data set was drawn from Florida's Birth Vital Statistics, 1997–1998. This data set contained sociodemographic and perinatal health factors, as well as a measure of tobacco use, for pregnant women who had children born in the state of Florida in 1997 and 1998. It was augmented by three other data sources: (1) Florida Birth Defects Registry, which contained information on children who were diagnosed as having BD; (2) Children's Medical Services Early Intervention Program, which contained information on children who were assessed for DDD and received evaluation or intervention services, 1997–2001; and (3) Medicaid Status from the Agency for Health Care Administration (AHCA) Medicaid eligibility files. Two exclusion criteria were applied to this merged population-based data set: (1) multiple births; (2) missing values of birth weight (BW), demographic, behavioral or perinatal health factors. The first criterion was used to satisfy the independency of individuals in the study population. After applying the above criteria and deleting $BW < 350$ and $BW \geq 6000$, 385 485 records were available for analysis. In this data set, dichotomous outcomes were BD diagnosed in the first year of life, AC of the newborn and DDD diagnosed under the age of 1. BW was classified into four categories: extremely LBW (ELBW, 350–999 g), very LBW (VLBW, 1000–1499 g), LBW (LBW, 1500–2499 g) and normal BW (NBW, 2500–5999 g). Health-related and sociodemographic factors and prenatal tobacco use were studied to assess their effects on each outcome (BD, AC, DDD and BW) and their effects on the IMI. Five health-related factors were previous pregnancy failure (PREVFAIL: Yes, No; Yes if one or more previous pregnancy terminated in either a spontaneous or induced abortion or if one or more previous pregnancy resulted in a liveborn infant who later died, and No if otherwise), 12 months or greater previous pregnancy interval (PREV12: Yes, No; Yes if pregnancy interval more than or equal to 12 months, and No if otherwise), experience of previous pregnancy success (PREVSOME: Yes, No; Yes if one or more than one previous pregnancies resulted in liveborns who were still living and no previous failures, and No if otherwise), healthy start services (HSSCREEN: Yes, No; Yes if one of the following services was delivered: administrative contact, care coordination or direct Healthy Start

Services such as counseling, education and other support, and No if no contact at all) and Florida's healthy start prenatal risk screen score (HSSCORE: numerical variable scored from 1 to 8 with higher score indicating higher risk). Six sociodemographic characteristics included in this study were mother's education level (MEDU), mother's marital status (MSTAT), poverty (MEDICAID), mother's age (MAGE), infant's race (BRACE) and infant's sex (BSEX). Mother's marital status (MSTAT: Yes, No; Yes if married and No if not married), poverty (MEDICAID: Yes, No; Yes if receiving medicaid served and No if otherwise) and infant's gender (BSEX: Yes, No; Yes for male and No for female) were dichotomous. Mother's education (MEDU) had three categories: less than high school (<HS, if mother's education was less than or equal to 11 years), high school (HS, if mother's education was equal to 12 years) or greater than high school (>HS, if mother's education was greater than 12 years). Mother's age (MAGE) also had three categories: <20, 20–34 and >34 years old. There were four categories for infant's race (BRACE): Black, White, Hispanic and Others. One more maternal variable, prenatal tobacco use (SMOKE), was continuous and was scored by the number of cigarettes per day by mothers in pregnancy.

3. STATISTICAL MODELS

3.1. Extended latent variable models

To specify extended LV models for multiple multinomial morbidity outcomes with complete responses, we define S to be a univariate unobservable LV of interest with values $s \in [0, \infty)$ and Y_m ($m = 1, \dots, M$) to be the m th manifestation of S with potential values $y_m \in \{1, \dots, C_m\}$. For simplicity, here we omit the subscript denoting individuals when giving the model for a single individual. Denote by θ_s the marginal density of S and by $\theta_{ms}^{y_m} = \pi(Y_m = y_m | s)$ the conditional probability that given s , the individual will have a response y_m to outcome m ($m = 1, \dots, M$; $y_m = 1, \dots, C_m$). Under LV models, all these outcome variables are associated because the population under study is a scaled mixture of subpopulations. As applied to our infant morbidity data, four manifestations of S are: Y_1 , BD with two categories (Yes, No); Y_2 , AC with two categories (Yes, No); Y_3 , DDD that is dichotomous (Yes, No); Y_4 , BW, which has four categories (ELBW, VLBW, LBW and NBW). In this case, it is natural to associate the LV S with the underlying infant morbidity and consider observed outcomes to be surrogates for S , which can be considered to be the propensity for infant morbidity. Here, θ_s denotes the marginal distribution of propensity for infant morbidity, and $\theta_{ms}^{y_m}$ is the probability of individuals who will have the response y_m to infant morbidity outcome m ($m = 1, 2, 3, 4$) given latent value s . Since S is the propensity for infant morbidity, it is reasonable to treat it as a continuous variable. Then, the joint distribution of morbidity outcomes can be written as

$$P(Y_1 = y_1, \dots, Y_M = y_M) = \int_0^\infty \theta_s \prod_{m=1}^M \prod_{j=1}^{C_m} (\theta_{ms}^j)^{y_{mj}} ds \quad (1)$$

where $y_{mj} = 1$ if the individual falls into the j th category of outcome m and 0 otherwise. The expression in (1) will be referred to as the cell probability in Section 4, which is the function of the vector of $\theta_{ms}^{y_m}$ and θ_s . The term $\prod_{j=1}^{C_m} (\theta_{ms}^j)^{y_{mj}}$ is a multinomial process associated with outcome m , where θ_{ms}^j is the conditional probability for the j th category of outcome m subject to $\sum_{j=1}^{C_m} \theta_{ms}^j = 1$ ($m = 1, \dots, M$). For BD, AC and DDD, it is a Bernoulli, but for BW, it is a

multinomial distribution with $n = 1$. Note that in (1), the LV is continuous and observed outcomes are categorical. In fact, the distribution (1) is a scaled mixture of product multinomial processes with mixing weights θ_s . One basic assumption implicit in (1) is that given the latent value s , responses of morbidity outcomes are independent, i.e.

$$P(Y_1 = y_1, \dots, Y_M = y_M | s) = \prod_{m=1}^M \prod_{j=1}^{C_m} (\theta_{ms}^j)^{y_{mj}} \quad (2)$$

This conditional independence is equivalent to the axiom of local independence [24, 25]. In our analysis, as discussed by Roche *et al.* [26], not only it is convenient, but also it defines the sense in which the S serves as a summary construct. Information about S is available from (1) through posterior distributions of the LV:

$$f(S | Y_1 = y_1, \dots, Y_M = y_M) = \frac{\theta_s \prod_{m=1}^M \prod_{j=1}^{C_m} (\theta_{ms}^j)^{y_{mj}}}{\int_0^\infty \theta_t \prod_{m=1}^M \prod_{j=1}^{C_m} (\theta_{mt}^j)^{y_{mj}} dt} \quad (3)$$

These posterior distributions capture information about the unknown LV given observed indicators and hence are useful in the development of a composite morbidity index for our example.

3.2. Modelling conditional distributions

Suppose that data of M distinct morbidity outcomes with C_m ($m = 1, \dots, M$) categories in outcome m are collected on an infant. Let $\theta_{ms}^{y_m} = f(Y_m = y_m | S = s)$ for $y_m = 1, \dots, C_m$ and $s \in [0, \infty)$ be the conditional probability that the infant will have a response y_m to outcome m . We capture information to develop a composite morbidity index through modeling the conditional probability of each morbidity outcome as a function of LV S . Since S is the continuous LV associated with manifest morbidity outcomes with domain $[0, \infty)$, it is reasonable to assume that the conditional probability of an infant having an adverse morbidity outcome will be zero when S is equal to zero. For this reason, the following models for conditional probability are considered:

$$\theta_{ms}^{y_m} = \frac{(\beta_{my_m} s)^{2\alpha_{my_m}}}{1 + \sum_{j=1}^{C_m-1} (\beta_{mj} s)^{2\alpha_{mj}}} \quad (4)$$

where β_{my_m} and α_{my_m} ($m = 1, \dots, M$; $y_m = 1, \dots, C_m - 1$) are parameters linking the LV variable to the conditional probability $\theta_{ms}^{y_m}$ subject to $\sum_{y_m=1}^{C_m} \theta_{ms}^{y_m} = 1$. They are the parameters of interest we need to estimate in the conditional distribution of S given the observed outcomes.

For every morbidity outcome, we treat the normal category as the reference category. That is to say, we specify $\alpha_{mC_m} = 0$ for any $m = 1, \dots, M$. For example, BD is the first dichotomous outcome (Yes, No) we introduced in our study. So the category BD = No is the reference category ($\alpha_{12} = 0$) and the model for the conditional probability of BD = Yes is

$$\theta_{1s}^1 = \frac{(\beta_{11} s)^{2\alpha_{11}}}{1 + (\beta_{11} s)^{2\alpha_{11}}}$$

BW is the fourth outcome under study with four categories (ELBW, VLBW, LBW and NBW). Then, the model for the conditional probability of ELBW is

$$\theta_{4s}^1 = \frac{(\beta_{41}s)^{2\alpha_{41}}}{1 + \sum_{j=1}^3 (\beta_{4j}s)^{2\alpha_{4j}}}$$

Similarly, we can easily write out the models for VLBW and LBW. Here, we treat NBW as the reference category ($\alpha_{44} = 0$).

3.3. Distribution of the latent variable

In our study, the LV S represents the propensity for infant morbidity which is manifest in four outcomes: BD, AC, DDD and LBW. Our prior belief is that most infants are exposed to low-level risks such that the LV may have an asymmetric distribution with a long right tail. We assume that S is distributed as log-normal with density

$$\theta_s = \frac{1}{s\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\log(s) - \mu)^2\right\} \quad (5)$$

where μ and σ are the mean and standard deviation of the log-LV. For simplicity, we refer to (4) combined with (1) and (5) as s -models.

Model (4) has good properties similar to those of logistic models, with the difference that the conditional probabilities in model (4) will be zero if s is anchored zero. This is generally consistent with the hypotheses in studies of infant medical and health care that prevalence for some disease is zero in the absence of risk factors. The relationship between (4) and the logistic model is reflected in the following reparameterization:

$$\theta_m^{y_m} = \frac{\exp(\beta_{0m}^* + \beta_{my_m}^* z)}{1 + \sum_{j=1}^{C_m-1} \exp(\beta_{0j}^* + \beta_{jy_m}^* z)} \quad (6)$$

where $\beta_{0m}^* = 2\alpha_{my_m}(\mu + \log \beta_{my_m})$, $\beta_{my_m}^* = 2\alpha_{my_m}/\sigma$ and $z = (\log(s) - \mu)/\sigma$ ($m = 1, \dots, M$; $y_m = 1, \dots, C_m$). Here, z is the value of the variable Z , which has a standard normal distribution. The relationship of (4) to the logistic model, together with the medical hypothesis, provides us a good reason to use these models in our analysis. Note that we require $s > 0$ and $\beta_{my_m} > 0$ in model (4) to complete this translation.

4. ESTIMATION METHOD

Consider M infant morbidity outcomes with each having C_m categories. Define β to be the vector of β_{my_m} and α_{my_m} ($m = 1, \dots, M$; $y_m = 1, \dots, C_m - 1$), which are parameters of interest in s -models defined in (4). Suppose that observations on all subjects can be arranged in $N = \prod_{m=1}^M C_m$ cell counts. Let $n = (n_1, \dots, n_N)$ be the vector of cells, where n_i 's ($i = 1, \dots, N$) have a multinomial distribution with the vector of cell probabilities $\pi(\beta) = (\pi_1(\beta), \dots, \pi_N(\beta))'$. Here, $\pi(\beta)$ is defined by expressing (1) and (4) as functions of $\beta = (\alpha_{11}, \beta_{11}, \dots, \alpha_{M(C_M-1)}, \beta_{M(C_M-1)}, \mu, \sigma)$. Let $p = (p_1, \dots, p_N)'$ be the vector of sample proportions ($p_i = n_i/n$). Then, sample proportions

are unbiased and consistent estimates of cell probabilities π . That is,

$$p = \pi(\beta) + \varepsilon$$

where the expectation of ε is $E(\varepsilon) = 0$ and the variance is $\text{var}(\varepsilon) = n^{-1}V$. From (1) and (4), it is easy to see that $\pi(\beta)$ is a nonlinear function of β . Following the nonlinear least-square theory [27], the estimated generalized nonlinear least square estimates of β can be obtained by minimizing the quadratic form:

$$Q(\beta, \hat{V}) = n(p - \pi(\beta))' \hat{V}^{-1} (p - \pi(\beta)) \quad (7)$$

where \hat{V} represents a consistent estimator of V and hence does not depend on β . In our study, we choose $\hat{V} = \text{diag}(p) - pp'$, which depends only on the sample proportion p . By taking the derivative on both sides of (7), we obtain the estimating equations

$$\frac{\partial Q(\beta, \hat{V})}{\partial \beta} = \left(\frac{\partial \pi(\beta)}{\partial \beta} \right)' \hat{V}^{-1} (p - \pi(\beta)) = 0 \quad (8)$$

Note that if we substitute V for \hat{V} , (8) becomes the optimal estimating equations which are used to compute the maximum-likelihood estimates (MLEs). Because of the dependence of V on β , the MLEs often attain local maxima. The idea of using the simplification (8) obtained by substituting \hat{V} for V in the optimal estimating equations is to avoid the local maxima problem that we face in computing the MLEs. Let $\beta^{(k)}$ denote the estimate at the k th iteration. By Hartley's modified Gauss-Newton method [28], the estimate $\beta^{(k+1)}$ at iteration $k + 1$ is

$$\beta^{(k+1)} = \beta^{(k)} + \lambda^{(k)} \left[\left(\frac{\partial \pi(\beta^{(k)})}{\partial \beta} \right)' \hat{V}^{-1} \left(\frac{\partial \pi(\beta^{(k)})}{\partial \beta} \right) \right]^{-1} \left(\frac{\partial \pi(\beta^{(k)})}{\partial \beta} \right)' \hat{V}^{-1} (p - \pi(\beta^{(k)})) \quad (9)$$

where $\lambda^{(k)} \in [0, 1]$ is the stepping coefficient for the k th iteration. We adjust $\lambda^{(k)}$ to guarantee that $\beta^{(k+1)}$ is a better approximation to the least-square estimator $\hat{\beta}$ than $\beta^{(k)}$ in the sense that $Q(\beta^{(k+1)}, \hat{V}) \leq Q(\beta^{(k)}, \hat{V})$. Although $\pi(\beta)$ and the derivatives of $\pi(\beta)$ with respect to β have no closed forms, we can calculate these integrals using common numerical integration methods since the integral over S is only one dimensional. In our study, there are four morbidity outcomes: BD, AC, DDD and BW. The first three are dichotomous and the last one is polytomous with four categories. Based on s -models, we have 13 parameters to be estimated. Note that considering the identifiability of parameters in s -models, we fix the variance parameter σ in (5) at 1 in our estimation.

5. ESTIMATION RESULTS

5.1. Calculation of initial estimates

To make (9) converge faster, selection of appropriate initial values is necessary. There are six independent pieces of information in the raw data that we can utilize to calculate initial values of parameters in s -models: three are from BD, AC and DDD, respectively; and another three are from BW. They are the prevalences of adverse outcomes for each morbidity outcome. In Table I,

Table I. Occurring rates for each category of outcomes.

Variable	Category	Frequency	Per cent
BD	Yes	9459	2.44
	No	377 426	97.56
AC	Yes	25 903	6.70
	No	360 982	93.30
DDD	Yes	8108	2.10
	No	378 777	97.90
BW	350–999	2921	0.76
	1000–1499	2980	0.77
	1500–2499	25 071	6.48
	2500–5999	355 913	91.99

we list the frequencies and prevalences corresponding to the four outcomes. The prevalences of adverse outcomes associated with BD, AC and DDD are 2.44, 6.70, and 2.10 per cent, respectively. Those associated with BW are 0.76, 0.77, and, 6.48 per cent, which correspond to ELBW, VLBW and LBW, respectively. Based on these pieces of information, we derive initial values for s -model parameters as follows:

Step 1: Derive the initial estimate of μ . Denote μ_0 to be the initial estimate of μ . Since $\log(S)$ has a normal distribution with mean μ , let μ_0 be $\log(s_0)$, where s_0 is the approximate average risk level in which infants in the study currently are. The term s_0 can be calculated from (4) using prevalence of the adverse outcome for BD in Table I. The estimate of s_0 is 0.158, and hence μ_0 is -1.844 .

Step 2: Derive initial values of other parameters in s -models. In this step, we fix all other α 's to be 1 besides α_{11} and β_{11} . Since s_0 is the approximate average risk, the values of θ_{ms}^{ym} 's (from (4)) corresponding to s_0 can be treated as prevalences of adverse outcomes for associated morbidity outcomes, which are listed in Table I. Based on the information in Table I, initial values of parameters in s -models are easily deduced. For example, β_{21} is a parameter associated with AC. From Table I, the prevalence of AC is 6.7 per cent. Using model (4), the initial value of β_{21} can be calculated as $\beta_{21}^{(0)} = \sqrt{0.067/(1 - 0.067)}/s_0 \approx 1.69$. Similarly, we can get initial values of other parameters corresponding to s -models.

Using the derived set of initial values will make modified Gauss–Newton algorithms work more efficiently in the sense that convergence will take much shorter time.

5.2. Results

Using the modified Gauss–Newton algorithm described in Section 4 and the group of initial values derived in Section 5.1, we obtain estimates of parameters in s -models (Table II). Substituting these estimates into model (4), we calculate conditional probabilities given s for each outcome.

Figure 1 shows the conditional probabilities for each outcome variable and the density of infant morbidity propensity (last panel in Figure 1). Change patterns are similar for three dichotomous outcomes: BD, AC and DDD. The conditional probabilities for adverse outcomes (BD: Yes, AC Yes, DDD: Yes) decrease, and those for normal outcomes (BD: No, AC: No, DDD: No), increase with increasing s . In the graph for the multinomial outcome BW, there are four lines that correspond to

Table II. Estimates of parameters in s -models.

BD		AC		DDD		BW					
α_{11}	β_{11}	α_{21}	β_{21}	α_{31}	β_{31}	α_{41}	β_{41}	α_{42}	β_{42}	α_{43}	β_{43}
1.01	1.07	1.12	2.07	2.36	2.09	3.52	2.17	3.13	2.14	1.26	2.39

NBW, LBW, VLBW and ELBW, respectively. Note that the trends of curves associated with LBW and VLBW are different from those of NBW and ELBW. They increase first, attain maximum, and then decrease. The changes for curves of NBW and ELBW are similar to those of BD, AC and DDD. The order in which curve peaks for BW appear is NBW, LBW, VLBW and ELBW.

Based on (1) and (4), we calculate the expected probabilities and counts for each combination of four morbidity outcomes. Expected probabilities and counts, together with observed proportions and counts, are listed in Table III. Since the sample size in our study is very large ($n = 385\,485$), it will not be appropriate to use χ^2 statistics to test the goodness of fit of s -models [29]. We suggest using the weighted regression method to do it. To perform the test, assume that ε is a random vector with mean $E(\varepsilon) = 0$ and variance $V(\varepsilon) = n[\text{diag}(\pi) - \pi\pi']$. Let the sample count be the dependent variable, denoted by Y , and the expected count be the covariate, denoted by X . Then, we can use SAS PROC REG to fit the model

$$Y^* = X^*\beta + \varepsilon^*$$

where $Y^* = \hat{V}^{-1/2}Y$, $X^* = (\hat{V}^{-1/2}1, \hat{V}^{-1/2}X)$ and $\varepsilon^* = \hat{V}^{-1/2}\varepsilon$. In the above transformation, $\hat{V} = n[\text{diag}(p) - pp']$ is the consistent estimator of V and $\hat{V}^{-1/2}$ is equal to $ED^{-1/2}E'$, where E is the matrix of eigenvectors of \hat{V} and D is the diagonal matrix with eigenvalues of \hat{V} down the diagonal. The test result ($R^2 = 0.98$) shows that our s -models fit the data set well.

6. DEVELOPMENT AND VALIDATION OF IMI

6.1. Development of IMI

In Section 5, we derived estimates of parameters related to s -models. Substituting parameter estimates into (3), we obtained posterior distributions of LV S (propensity for infant morbidity). Let $\hat{s} = \hat{E}(S|Y = y)$ then,

$$\hat{s} = \int sf(s|Y = y) ds \quad (10)$$

where $f(s|Y = y)$ is given by (3), and Y is the vector of M manifestations: Y_1, \dots, Y_M .

To adapt for use in infant health, we re-scale \hat{s} on the 1–100 scale and arrive at the proposed IMI. Thus, the IMI is estimated as a function of its manifestations: BD, AC, DDD and BW. It is a composite index of propensity for infant morbidity which is developed from s -models. Table IV lists results for \hat{s} and IMI, given the pattern of four morbidity outcomes. We see that all infants who have the same pattern of outcomes have equal IMI values. The IMI can identify patterns

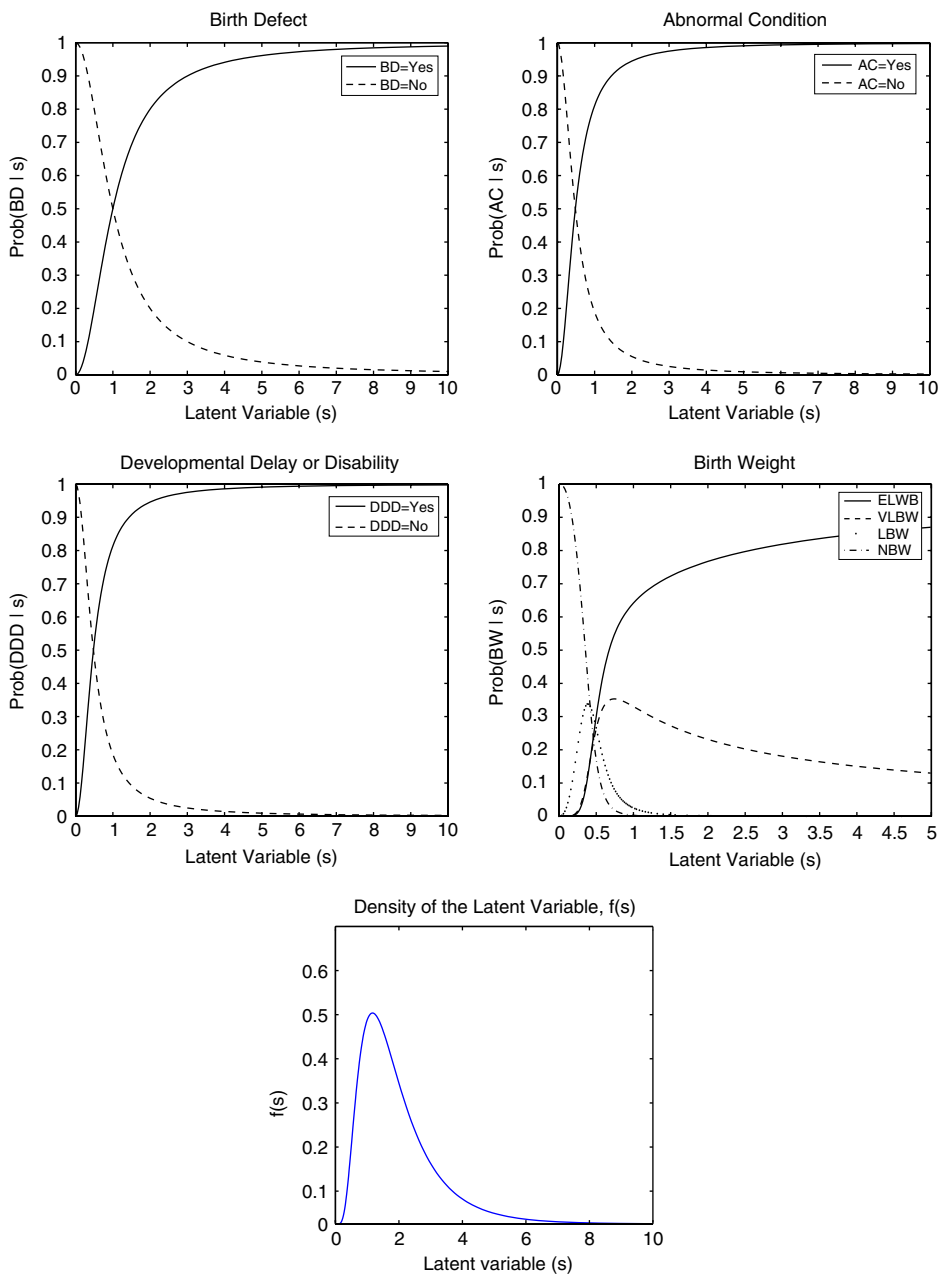


Figure 1. Curves of conditional probabilities for each outcome and density of the latent variable.

Table III. Observed counts (OBCNT), expected counts (EXCNT), observed per cents (OBPCNT) and expected per cents (EXPCNT) by combinations of four morbidity outcomes.

BD	AC	DDD	BW	OBCNT	EXCNT	OBPCNT (per cent)	EXPCNT (per cent)
Yes	Yes	Yes	350–999	94	123.80	0.0243	0.0320
Yes	Yes	Yes	1000–1499	86	100.52	0.0222	0.0260
Yes	Yes	Yes	1500–2499	151	98.71	0.0390	0.0255
Yes	Yes	Yes	2500–5999	278	96.79	0.0719	0.0250
Yes	Yes	No	350–999	68	92.08	0.0176	0.0238
Yes	Yes	No	1000–1499	84	88.92	0.0217	0.0230
Yes	Yes	No	1500–2499	262	263.66	0.0677	0.0682
Yes	Yes	No	2500–5999	747	731.35	0.1931	0.1890
Yes	No	Yes	350–999	87	97.14	0.0225	0.0251
Yes	No	Yes	1000–1499	101	84.41	0.0261	0.0218
Yes	No	Yes	1500–2499	339	130.59	0.0876	0.0338
Yes	No	Yes	2500–5999	929	195.63	0.2401	0.0506
Yes	No	No	350–999	75	125.71	0.0194	0.0325
Yes	No	No	1000–1499	128	137.01	0.0331	0.0354
Yes	No	No	1500–2499	707	861.22	0.1827	0.2226
Yes	No	No	2500–5999	5323	5312.77	1.3759	1.3732
No	Yes	Yes	350–999	471	422.99	0.1217	0.1093
No	Yes	Yes	1000–1499	319	364.50	0.0825	0.0942
No	Yes	Yes	1500–2499	441	534.50	0.1140	0.1382
No	Yes	Yes	2500–5999	581	762.32	0.1502	0.1970
No	Yes	No	350–999	778	512.65	0.2011	0.1325
No	Yes	No	1000–1499	671	550.99	0.1734	0.1424
No	Yes	No	1500–2499	3318	3167.57	0.8576	0.8187
No	Yes	No	2500–5999	17 554	17 684.24	4.5373	4.5709
No	No	Yes	350–999	429	403.16	0.1109	0.1042
No	No	Yes	1000–1499	447	380.44	0.1155	0.0983
No	No	Yes	1500–2499	994	980.62	0.2569	0.2535
No	No	Yes	2500–5999	2361	2370.33	0.6103	0.6127
No	No	No	350–999	919	977.03	0.2375	0.2525
No	No	No	1000–1499	1144	1217.11	0.2957	0.3146
No	No	No	1500–2499	18 859	18 901.28	4.8746	4.8855
No	No	No	2500–5999	328 140	329 114.95	84.8159	85.0679

of four outcomes. The range of the IMI is from 1 to 100. The maximal value is generated by infants who have the outcome pattern of BD = Yes, AC = Yes, DDD = Yes and ELBW = Yes (the first row), and the minimal value is associated with infants whose pattern is BD = No, AC = No, DDD = No and NBW = Yes (the last row). In addition, given the pattern of any three outcomes, there is an obvious trend of the IMI with the last outcome, i.e. values of the IMI increase with the last outcome getting worse. Further study will be needed to determine whether the IMI can exactly identify the pattern of outcomes in general situations.

6.2. Validation analysis

Before the IMI can be used in the medical/public health area, evaluation of the validity of the IMI is necessary. Since the IMI is a summary index of BD, AC, DDD and LBW, we expect that it will be associated with the likelihood of each morbidity outcome. To validate the index IMI, we need

Table IV. Estimates of posterior mean of the latent variable \hat{s} and the IMI.

BD	AC	DDD	BW	\hat{s}	IMI
Yes	Yes	Yes	350–999	0.5598	100.0000
Yes	Yes	Yes	1000–1499	0.5460	97.0435
Yes	Yes	Yes	1500–2499	0.4580	78.1904
Yes	Yes	Yes	2500–5999	0.3863	62.8295
Yes	Yes	No	350–999	0.4523	76.9693
Yes	Yes	No	1000–1499	0.4308	72.3631
Yes	Yes	No	1500–2499	0.3207	48.7754
Yes	Yes	No	2500–5999	0.2390	31.2720
Yes	No	Yes	350–999	0.5154	90.4878
Yes	No	Yes	1000–1499	0.4971	86.5672
Yes	No	Yes	1500–2499	0.3948	64.6505
Yes	No	Yes	2500–5999	0.3195	48.5183
Yes	No	No	350–999	0.3887	63.3436
Yes	No	No	1000–1499	0.3660	58.4804
Yes	No	No	1500–2499	0.2488	33.3716
Yes	No	No	2500–5999	0.1613	14.6256
No	Yes	Yes	350–999	0.5207	91.6232
No	Yes	Yes	1000–1499	0.5029	87.8098
No	Yes	Yes	1500–2499	0.4016	66.1073
No	Yes	Yes	2500–5999	0.3267	50.0608
No	Yes	No	350–999	0.3955	64.8005
No	Yes	No	1000–1499	0.3730	59.9801
No	Yes	No	1500–2499	0.2567	35.0641
No	Yes	No	2500–5999	0.1694	16.3610
No	No	Yes	350–999	0.4655	79.7972
No	No	Yes	1000–1499	0.4444	75.2768
No	No	Yes	1500–2499	0.3355	51.9461
No	No	Yes	2500–5999	0.2554	34.7855
No	No	No	350–999	0.3291	50.5750
No	No	No	1000–1499	0.3053	45.4761
No	No	No	1500–2499	0.1795	18.5248
No	No	No	2500–5999	0.0977	1.0000

to show that

- IMI is correlated with an alternative FVI defined below.
- IMI is associated with IM. In our data set, we define IM as a dichotomous outcome variable (Yes, No).
- IMI identifies categories of risk factors, i.e. least-square means (LSMs) of the IMI are different between/among categories of related factors.
- The effects of risk factors on IMI are representative of their effects on manifest morbidity outcomes.
- IMI discriminates morbidity outcome categories.

An FVI based on four manifestations is defined as follows. We count one point for each adverse outcome of the following events: BD(Yes|No), AC(Yes|No), DDD(Yes|No), LBW(Yes|No), VLBW(Yes|No) and ELBW(Yes|No). The FVI will have a score range from 0 to 6 by summing

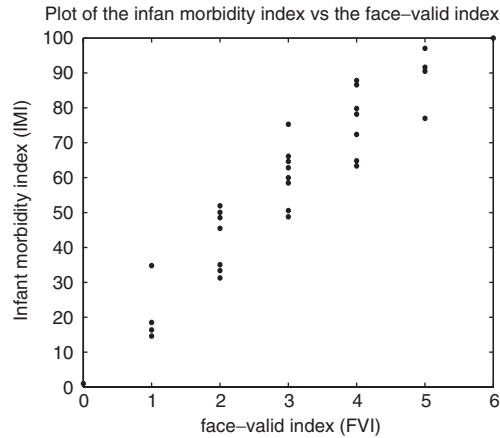


Figure 2. Scatterplots of the infant morbidity index (IMI) *versus* the face-valid index (FVI).

all points for each combination of these events. For example, if an infant falls into the cell which corresponds to 'yes' for BD and AC and 'No' for other events, then the FVI for this infant is equal to '2'; if 'No' occurs for all events, then the FVI is equal to '0'; similarly, if 'Yes' occurs for all events, then the FVI is '6'. This method for defining a new single variable which is a summary of given manifest outcomes not only provides useful comprehensive information, but also simplifies the procedure for data analysis. It is currently applied as a summary index in data analysis for multiple outcomes. In Section 6.1, we developed a new index IMI using LV models based on the same outcomes on which we defined the FVI. The question is 'how well does the IMI correlate with the FVI?'. To answer this question, we perform a simple regression analysis in which the IMI is the dependent variable and the FVI is the independent variable. From regression of the IMI on the FVI, we obtain the determination coefficient $R^2 = 0.94$, which means that in variation of the IMI, 94 per cent can be explained by that of the FVI. In addition, we draw a scatter plot of the FVI against the IMI (Figure 2). From this figure, we can also see the relationship between the IMI and the FVI. These analyses show that the IMI is highly related to the FVI. This aspect of validity of the IMI is referred to as 'concurrent validity' [30, 31].

As described in Section 1, BD, AC, DDD and LBW are four major outcomes that are associated with infant morbidity and IM. The mean value of IMI associated with infants who fall into the category of IM = Yes is 34.1 and that in the category of IM = No is 14.4. By a simple *t*-test, we have the *p*-value less than 0.0001, which means that there exists a significant difference in LSMs of the IMI between categories of IM and that the IMI can identify the pattern of IM as the individual outcome did. So the IMI is related to IM. This aspect of validity of the IMI is referred to as 'predictive validity' [31, 32]. We fit logistic regression models to compare the impact of the IMI, the sumscore FVI and the combination of manifest morbidity outcomes on IM. Measures of goodness of fit of predicted models are listed in Table V. From this table, we can see that the model with the IMI has the best goodness of fit for predicting IM (the smaller the criterion, the better the model). Hence, the new index IMI may have the combined impact of the four individual morbidity measures on childhood adverse outcomes such as IM.

To show (c) and (d), health-related and sociodemographic factors, and prenatal tobacco use are studied to assess their effects on each of BD, AC, DDD and BW and their effects on the IMI.

Table V. Measures of goodness of fit of predicted models for infant mortality.

Variable	AIC	SIC	-2 Log <i>L</i>
BD + AC + DDD + LBW	22 988.61	23 216.71	22 946.61
FVI	23 026.60	23 222.12	22 990.60
IMI	21 814.87	22 010.39	21 778.87

AIC, Akaike information criterion; SIC, Schwarz information criterion.
The models were adjusted by the same cluster of covariates used in Table VII.

Table VI. Least square means of the IMI by categories of risk factors.

PREVFAIL	Yes	No	
	13.9	12.7	
PREV12	Yes	No	
	13.1	14.2	
PREVSOME	Yes	No	
	12.8	14.7	
HSSCREEN	Yes	No	
	13.5	14.2	
BSEX	Male	Female	
	14.8	14.5	
MSTAT	Not married	Married	
	14.9	14.3	
MCAID	Yes	No	
	14.9	14.6	
MEDU	<HS	HS	>HS
	15.0	14.7	14.6
MAGE	>34	20–34	<20
	15.2	14.6	14.5
BRACE	Black	White	Others
	15.4	14.7	14.6
			Hispanic
			14.3

For convenience of analysis, we display LSMs of the IMI by each category of 10 categorical risk factors in Table VI. From this table, we can see that infants with adverse outcomes of risk factors have larger LSMs of the IMI than those with normal ones. Simple *t*-tests show that differences of LSMs of the IMI between/among categories of risk factors are all very significant ($p < 0.01$) except the difference between categories of White and others in BRACE ($p = 0.0462$) and the one between categories of <20 and 20–34 in MAGE ($p = 0.0156$). In fact, the latter two cases attain 0.05 significance level. Also, we analyze the effects of SMOKE and HSSCORE on the IMI. The IMI increases 1.6 with one more pack smoked. The effect of 2 packs/day is greater than any other effect. For each added stressor, the IMI increases by 0.41. The effect of stress (HSSCORE) is approximately $\frac{1}{2}$ that of <HS versus >HS and almost equals that of unmarried status. In the interest of brevity, we just demonstrate key results and describe others related to our purposes. Thus, we may come to a conclusion that the IMI can identify patterns of related categorical risk factors as well as each of the four morbidity outcomes. For convenience, we refer to this aspect of validity of the IMI as ‘identifiable validity’. This is a new validation procedure we propose in this study.

Table VII. Effects of risk factors on each outcome and the IMI.

Factor	Category	IMI	BD	AC	DDD	BW
PREVFAIL	Yes	0.45*	0.042*	0.113*	0.09*	0.04*
PREV12	Yes	-0.14*	-0.023	0.0095	-0.07*	-0.08*
PREVSOME	Yes	-0.19*	-0.044*	-0.072*	0.01	-0.01
HSSCREEN	Yes	-0.14*	-0.011	-0.077*	-0.02	-0.09*
BSEX	Male	0.06*	0.177*	0.066*	0.09*	-0.11*
MSTAT	Not married	-0.25*	-0.02	-0.0044	-0.07*	-0.08*
MEDICAID	Yes	0.29*	0.115*	0.067*	0.14*	0.06*
MEDU	<HS	0.35*	0.069*	0.093*	0.16*	0.06*
	HS	0.09*	0.0053	-0.014	-0.06*	-0.01
MAGE	<20	-0.08*	-0.192*	-0.091*	-0.25*	-0.06*
	>34	0.59*	0.215*	0.115*	0.32*	0.19*
BRACE	Black	0.7*	0.044*	0.056*	0.29*	0.33*
	Hispanic	-0.38*	-0.0045	-0.261*	-0.14*	-0.23*
	Others	-0.13*	-0.202*	0.007	-0.27*	0.06*
SMOKE		0.04*	0.0029*	0.0033*	0.01*	0.02*
HSSCORE		0.18*	0.054*	0.01	0.12*	0.12*

Note: Estimates with '*' are significant.

By GLM and logistic analysis, we demonstrate the effects of 12 risk factors on the IMI and each of individual outcomes in Table VII. We see that all risk factors have significant effects on the IMI, but their effects on BD, AC, DDD and BW are in different cases, with some effects significant and others not. For example, PREVFAIL, PREVSOME, Black and others of BRACE, BSEX, <HS of MEDU, MEDICAID, MAGE, SMOKE and HSSCORE have significant effects on BD, but other factors do not have. So are AC, DDD and BW.

Next, we will show that the effects of risk factors on the IMI are consistent with those on each of outcomes BD, AC, DDD and BW. What we are mainly concerned about are answers to two key questions: (1) If a risk factor has a consistent effect (positive or negative) on four morbidity outcomes (Case 1), will it also have a consistent effect on the IMI? (2) If a risk factor does not have a consistent effect on four morbidity indicators (Case 2), will the IMI still be able to identify it? Here, 'consistent' means that a risk factor has a similar effect on all outcomes by increasing or decreasing infant morbidity. If the answers to these two questions are both 'Yes', then (d) holds. For brevity, in the rest of this section, when we mention that a factor has an effect on some outcome variable or the IMI, we mean that it has a significant effect. Similarly, no effects mean no significant effects. From Table VII, we can see that, overall, risk factors have consistent effects on the IMI as those on manifest outcomes. For example, PREVFAIL increases the likelihoods of BD, AC, DDD and LBW; it also increases the likelihood of the IMI. Hence, its effects on manifest outcomes and the IMI are consistent. In cases of inconsistency, effects of risk factors on the IMI are not consistent only with BW. For example, 'male' of baby race (BRACE) has consistent positive effects on the IMI with those on BD, AC and DDD. Also, in these cases, the IMI can still identify the patterns of risk factors. We would discuss this feature in the final section. The above results show that the effects of risk factors on the IMI reflect the overall impressions of those on the individual outcome. Thus, the use of the IMI as a single variable should not lead to misleading interpretations of effects on morbidity outcomes. Based on our analysis, it should be valid that the effects of risk factors on the IMI are representative of their effects on each of

the manifest morbidity outcomes. We refer to this aspect of validity of the IMI as 'consistent or representative validity', which is a new term proposed in the paper.

By simple regression of the IMI on each morbidity outcome, we arrive at determination coefficient R^2 between the IMI and BD, AC, DDD and BW as 0.13, 0.28, 0.43, and 0.31, respectively. Thus, the IMI is positively related to each infant morbidity outcome. Simple tests show that the differences of LSM's between/among categories of each outcome are significant. This means that the IMI can identify the patterns of the individual outcome. We may conclude that the IMI is related to four morbidity outcomes and can discriminate their patterns, i.e. (e) holds. This aspect of validity of the IMI is referred to as 'discriminative validity' [30, 33].

Hence, the IMI is correlated with an FVI of morbidity outcomes, IM and each of the individual outcome, and can be used in future research as a measure of infant propensity for morbidity.

7. DISCUSSION

LV models have been applied in many clinical settings to describe disease in populations. Recent applications include gerontology [26], genetics [34], medical care [35, 36], ophthalmology [37] and cytometry [38]. Specifically in psychiatric research, LV models have been used in studies of alcoholism [39–41], autism [42, 43], social phobias [44], schizophrenia [45], psychiatric syndromes [46, 47] and psychiatric disorder [48–51]. In this study, LV models were applied to the study of infant morbidity, in which we developed a composite IMI based on four major outcomes BD, AC, DDD and BW. This method for developing a composite single variable may be applied and extended to other areas.

In Section 6.2, we show that the IMI is related to four manifest outcomes and is a summary index. The IMI is highly correlated with a FVI and predicts IM. Also, it is predicted by risk factors in representative and expected ways. The IMI can be used in future work as a single composite morbidity outcome. Since it is a summary index of all outcome variables of interest, we can use it to capture more information than each outcome variable does to identify and quantify the effects of risk factors associated with infant morbidity. Compared with the FVI, the IMI allows more variation among observations and can identify all the patterns of morbidity outcomes. The single-variable analysis will not only provide a useful tool for accurate evaluation of risk factors but also simplify the procedure for data analysis. A similar procedure can be applied to validate the use of indices associated with other outcomes. We need to note that the validation procedure in Section 6.2 is necessary.

Our study demonstrates an approach to development and validation of a composite index where the LV is continuous and the observed outcomes are dichotomous or multinomial using LV modeling and modified Gauss–Newton estimation. In fact, this approach would have been latent trait setting since the LV S under study is continuous (we assumed that S was distributed as log-normal). This approach can be generalized to other LV models, such as latent profile models, where the LV is categorical and the observed indicators are continuous, and latent structure models, where the LV and the observed indicators are all continuous. In all these LV models, the key is to associate general health status such as infant morbidity, physical disability and mental illness status, of which there is no single measure analogous to, say, blood pressure as a measure of cardiac output, with the unobservable LV.

In validation analysis of the IMI, we find that the IMI can discriminate the pattern of each outcome: BD, AC, DDD and BW. In addition, the IMI can identify categories of risk factors in

the sense that LSMs of the IMI are different among categories of these factors. When we show the consistency of effects of risk factors on the IMI with those on manifest variables, a problem may occur if the effects of risk factors on manifest outcomes are not consistent. For example, in Table VII, the 'HS' of risk factor MEDU does not have consistent effects on DDD and BW. It lowers the likelihood of DDD, but increases the likelihood of LBW. In this case, it is not possible that the effect of 'HS' of MEDU on the IMI is consistent with effects on both DDD and BW. Although we cannot use this case to show the effect consistency of the IMI and manifest outcomes, we demonstrate the usefulness of the IMI in the sense that we may use the IMI to identify a risk factor when we cannot identify it by analysis of separate manifest outcomes.

In Table IV, estimates of the IMI are different and have regular changes associated with different combinations of four basic outcomes. For example, the infants who are observed to have adverse outcomes in the worst situation have the maximal value of the IMI, and the minimal value of the IMI corresponds to the infants who fall into the pattern of best situation, i.e. the pattern of not having adverse outcomes. It is obvious that the IMI can identify these two extreme patterns. But whether the IMI can exactly identify all combined patterns of the observed indicators still needs further study by repeated sample analysis or development of a procedure for testing the difference of the IMI among different patterns of observed outcome variables. On attaining this goal, the methodology demonstrated in this study will become widely available for use in the medical and public health areas.

ACKNOWLEDGEMENTS

We would like to thank the editor, one associate editor and two referees for their constructive comments and suggestions to improve this manuscript. Special thanks to Randolph Carter in the Department of Biostatistics at the New York State University, Buffalo, for his suggestions on model development and P. V. Rao in the MCHERDC at the University of Florida, Gainesville, for his comments on comparison of predicted models. We wish to acknowledge the following groups for their support:

1. Maternal Child Health and Education Research and Data Center, College of Medicine, University of Florida.
2. Children's Medical Services, Florida Department of Health.
3. Office of Medicaid Research and Policy, Florida Agency for Health Care Administration.
4. Lawton and Rhea Chiles Center for Healthy Mothers and Babies, College of Public Health, University of South Florida.

REFERENCES

1. O'Shea TM, Klinepeter KL, Goldstein DJ, Jackson BW, Dillard RG. Survival and developmental disability in infants with birth weights of 501 to 800 grams, born between 1979 and 1994. *Pediatrics* 1997; **100**:982–986.
2. Hack M, Klein NK, Taylor HG. Long-term developmental outcome of low birth weight infants. *The Future of Children* 1995; **5**:176–196.
3. Kirby RS. Co-occurrence of developmental disabilities with birth defects. *Mental Retardation and Developmental Disabilities Research Review* 2002; **8**:182–187.
4. Wattendorf DJ, Muenke M. Fetal alcohol spectrum disorders. *American Family Physician* 2005; **72**:279–282.
5. Astley SJ, Clarren SK. A case definition and photographic screening tool for the facial phenotype of fetal alcohol syndrome. *Journal of Pediatrics* 1996; **119**:33–41.
6. Agustines LA, Kub YG, Rumney PJ, Lu MC, Bonebrake R, Asrat T, Nageotte M. Outcomes of extremely low-birth-weight infants between 500 and 750 g. *American Journal of Obstetrics and Gynecology* 2000; **182**:1113–1116.
7. Hogan DP, Park JM. Family factors and social support in the developmental outcomes of very low-birth weight children. *Clinical Perinatology* 2000; **27**:433–459.

8. Schendel DE, Stockbauer JW, Hoffman HJ, Herman AA, Berg CJ, Schramm WF. Relation between very low birth weight and developmental delay among preschool children without disabilities. *American Journal of Epidemiology* 1997; **146**:740–749.
9. Hack M, Fanaroff AA. Outcomes of children of extremely low birth weight and gestational age in the 1990's. *Early Human Development* 1999; **53**:193–218.
10. Bucher HU, Killer C, Ochsner Y, Vaihinger S, Fauchere J. Growth, developmental milestones and health problems in the first 2 years in very preterm infants compared with term infants: population based study. *European Journal of Pediatrics* 2002; **161**:151–156.
11. Collin MF, Jalsey CL, Anderson CL. Emerging developmental sequelae in the 'normal' extremely low birth weight infants. *Pediatrics* 1991; **88**:115–120.
12. Saigal S, Stoskopf BL, Streiner DL, Burrows E. Physical growth and current health status of infants who were of extremely low birth weight and controls at adolescence. *Pediatrics* 2001; **108**:407–415.
13. Mathews TJ, Curtin SC, MacDorman MF. Infant mortality statistics from the 1998 period linked birth/infant death data set. *National Vital Statistics Reports* 2000; **48**:1–25.
14. Mathews TJ, MacDorman MF, Menacker F. Infant mortality statistics from the 1999 period linked birth/infant death data set. *National Vital Statistics Reports* 2002; **50**:1–28.
15. Vinceti M, Rovesti S, Bergomi M, Calzolari E, Candela S, Campagna A, Milan M, Vivoli G. Risk of birth defects in a population exposed to environmental lead pollution. *The Science of the Total Environment* 2001; **278**:23–30.
16. Shaw GM, Lammer EJ. Maternal periconceptional alcohol consumption and risk for orofacial clefts. *Journal of Pediatrics* 1999; **134**:298–303.
17. Lief S, Olshan AF, Werler M, Strauss RP, Smith J, Mitchell A. Maternal cigarette smoking during pregnancy and risk of oral clefts in newborns. *American Journal of Epidemiology* 1999; **150**:683–694.
18. Lorente C, Cordier S, Goujard J, Ayme S, Bianchi F, Calzolari E, De Walle HE, Knill-Jones R. Tobacco and alcohol use during pregnancy and risk of oral clefts. Occupational Exposure and Congenital Malformation Working Group. *American Journal of Public Health* 2000; **90**:420–423.
19. Thompson JR, Carter RL, Edwards AR, Roth J, Ariet M, Ross NL, Resnick MB. A population based study of the effects of birth weight on early developmental delay and disability in children. *American Journal of Perinatology* 2003; **20**:321–332.
20. Goldenberg RL, Hauth JC, Andrews WW. Intrauterine infection and preterm delivery. *New England Journal of Medicine* 2000; **342**:1500–1507.
21. Garrett ES, Eaton WW, Zeger SL. Methods for evaluating the performance of diagnostic tests in the absence of a gold standard: a latent class model approach. *Statistics in Medicine* 2002; **21**:1289–1307.
22. Lin HQ, McCulloch CE, Turnbull BW, Slate EH, Clark CL. A latent class mixed model for analysing biomarker trajectories with irregularly scheduled observations. *Statistics in Medicine* 2000; **19**:1303–1318.
23. Qu YS, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics* 1996; **52**:797–810.
24. McCutcheon AL. *Latent Class Analysis*. Sage: Newbury Park, CA, 1987.
25. Lazarsfeld PF, Henry NW. *Latent Structure Analysis*. Houghton-Mifflin: New York, 1968.
26. Roche KB, Miglioretti DL, Zeger SL, Rathouz PJ. Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association* 1997; **92**:1375–1386.
27. Gallant RA. *Nonlinear Statistical Model*. Wiley: New York, 1987.
28. Hartley HO. The modified Gauss–Newton method for the fitting of non-linear regression functions by least squares. *Technometrics* 1961; **3**:269–280.
29. Agresti A. *Categorical Data Analysis* (2nd edn). Wiley: New York, 2002.
30. Saigal S, Rosenbaum P, Stoskopf B, Hoult L, Furlong W, Feeny D, Hagan R. Development, reliability and validity of a new measure of overall health for pre-school children. *Quality Life Research* 2005; **14**:243–257.
31. Voigt RG, Brown FR, Fraley JK, Liorente AM, Rozelle J, Jensen CL, Heird WC. Concurrent and predictive validity of the cognitive adaptive test/clinical linguistic and auditory milestone scale (cat/clams) and the mental developmental index of the Bayley scales of infant development. *Clinical Pediatrics* 2003; **42**:427–432.
32. McKeith IG, Fairbairn AF, Bothwell RA, Moore PB, Ferrier IN, Thompson P, Perry RH. An evaluation of the predictive validity and inter-rater reliability of clinical diagnostic criteria for senile dementia of Lewy body type. *Neurology* 1994; **44**:872–877.
33. Benning SD, Patrick CJ, Salekin RT, Leistico AMR. Convergent and discriminant validity of psychopathy factors assessed via self-report. *Assessment* 2005; **12**:270–289.

34. Eaves LJ, Silberg JL, Hewitt JK, Rutter M, Meyer JM, Neale MC, Pickles A. Analyzing twin resemblance in multisympptom data: genetic applications of a latent class model for symptoms of conduct disorder in juvenile boys. *Behavioral Genetics* 1993; **23**:5–19.
35. Forbes JF, Pickering RM. Development of a neonatal case-mix classification. *Medical Care* 1988; **26**:1033–1045.
36. Baker D, Taylor H. Inequality in health and health service use for mothers of young children in South West England Survey team of the Avon Longitudinal Study of Pregnancy and Children Team. *Journal of Epidemiology and Community Health* 1997; **51**:74–79.
37. Bandeen-Roche K, Munoz B, Tielsch JM, West SK, Schein OD. Self-reported assessment of dry eye in a population-based setting. *Investigative Ophthalmology and Visual Science* 1997; **38**:2469–2475.
38. van Putten WL, de Vries W, Reinders P, Levering W, van der Linden R, Tanke HJ, Bolhuis RL, Gratama JW. Quantification of fluorescence properties of lymphocytes in peripheral blood mononuclear cell suspensions using a latent class model. *Cytometry* 1993; **14**:86–96.
39. Bucholz KK, Heath AC, Reich T, Hesselbrock VM, Kramer JR, Nurnberger Jr JI, Schuckit MA. Can we subtype alcoholism. *Alcoholism: Clinical and Experimental Research* 1996; **20**:1462–1471.
40. Kendler KS, Karkowski LM, Prescott CA, Pedersen NL. Latent class analysis of temperance board registrations in Swedish male–male twin pairs born 1902 to 1949: searching for subtype of alcoholism. *Psychological Medicine* 1998; **28**:803–813.
41. Fergusson DM, Horwood LJ, Lynskey MT. The prevalence and risk factors associated with abusive or hazardous alcohol consumption in 16-year-olds. *Addiction* 1995; **90**:935–946.
42. Szatmari P, Volkmar F, Walter S. Evaluation of diagnostic criteria for autism using latent class models. *Journal of the Academy of Child and Adolescent Psychiatry* 1995; **34**:216–222.
43. Pickles A, Bolton P, MacDonald H, Bailey A, LeCouteur A, Sim CH, Rutter M. Latent class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. *American Journal of Human Genetics* 1995; **57**:717–726.
44. Kessler RC, Stein MB, Berglund P. Social phobia subtypes in the National Comorbidity Survey. *American Journal of Psychiatry* 1998; **155**:613–619.
45. Sham PC, Castle DJ, Wessely S, Farmer AE, Murray RM. Further exploration of latent class typology of schizophrenia. *Schizophrenia Research* 1996; **20**:105–115.
46. Sullivan PF, Kendler KS. Typology of common psychiatric syndromes: an empirical study. *British Journal of Psychiatry* 1998; **173**:312–319.
47. Kendler KS, Karkowski LM, Walsh D. The structure of psychosis: latent class analysis of probands from the Roscommon Family Study. *Archives of General Psychiatry* 1998; **55**:492–499.
48. Melton B, Liang KY, Pulver AE. Extended latent class approach of the study of familial/sporadic forms of disease: its application of the study of heterogeneity of schizophrenia. *Genetic Epidemiology* 1994; **11**:311–327.
49. Eaton WW, Dryman A, Sorenson A, McCutcheon A. DSM-III major depressive disorder in community: a latent class analysis of data from the NIMH Epidemiologic Catchment Area Program. *British Journal of Psychiatry* 1989; **155**:48–54.
50. Nestadt G, Hanfelt J, Liang KY, Lamacz M, Wolyniec A, Pulver AE. An evaluation of the structure of schizophrenia spectrum personality disorders. *Journal of Personality Disorders* 1994; **8**:288–298.
51. Eaton WW, McCutcheon A, Dryman A, Sorenson A. Latent class analysis of anxiety and depression. *Sociological Methods and Research* 1989; **18**:104–125.