CSP CADERNOS DE SAÚDE PÚBLICA
REPORTS IN PUBLIC HEALTH

Factors associated with preterm birth: from logistic regression to structural equation modeling

Fatores associados ao nascimento pré-termo: da regressão logística à modelagem com equações estruturais

Factores asociados con el nacimiento prematuro: de la regresión logística al modelo de ecuaciones estructurales

Adelaide Alves de Oliveira [1]
Marcia Furquim de Almeida [1]
Zilda Pereira da Silva [1]
Paula Lisiane de Assunção [2]
Ana Maria Rigo Silva [3]
Hellen Geremias dos Santos [1]
Gizelton Pereira Alencar [1]

doi: 10.1590/0102-311X00211917

Abstract

This study proposed the application of structural equation modeling (SEM) to investigate variables associated with preterm birth based on a theoretical model analyzed previously by hierarchical logistic regression. The data came from a population-based case-control observational study of hospital births to mothers residing in Londrina, Paraná State, Brazil (June 2006 to March 2007). For the SEM, the study considered the association between socioeconomic characteristics and psychosocial aspects pertaining to reproductive history, work and physical activity, complications during the pregnancy, and fetal characteristics. It also considered the relationship between these associations and the outcome preterm birth mediated by adequacy of prenatal care. The weighted least square mean and variance adjusted estimator (WLSMV) was used for categorical data and robust maximum likelihood (MLR) for odds ratios. Three latent variables were created: socioeconomic vulnerability, family vulnerability, and unwanted pregnancy. The effect of socioeconomic and family vulnerability and unwanted pregnancy on prematurity occurred indirectly through inadequacy of prenatal care. The proposed methodology allowed using constructs, verifying the role of mediation by inadequacy of prenatal care, and identifying the variables' direct and indirect effects on the outcome preterm birth.

Statistical Models; Logistic Models; Risk Factors; Premature Birth

Correspondence
G. P. Alencar
Faculdade de Saúde Pública, Universidade de São Paulo.
Av. Dr. Arnaldo 715, São Paulo, SP 01246-904, Brasil.
gizelton@usp.br

[1] Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brasil.
[2] Departamento de Saúde de Jequié, Universidade Estadual do Sudoeste da Bahia, Jequié, Brasil.
[3] Centro de Ciências da Saúde, Universidade Estadual de Londrina, Londrina, Brasil.

Introduction

Theoretical models for studying preterm birth have demonstrated the importance of multiple variables associated with preterm birth. Gestational complications, including bleeding, high blood pressure, eclampsia, altered amniotic fluid volume, genital tract infection, and diabetes have a direct effect on gestational age at birth [1]. Various authors have pointed to multiple gestation as a risk factor for preterm birth, including the presence of complications, which may be related to the occurrence of premature rupture of membranes or other maternal-fetal disorders [1,2,3,4]. Maternal age, reproductive history including parity, and characteristics of previous children (low birth weight and prematurity) have been identified by various studies as risk factors for preterm birth [1,5,6,7]. Aspects of inadequate prenatal care, such as late initiation (third trimester) and unsatisfactory frequency (fewer than 6 consultations) are also risk factors for preterm birth [5,6,8].

Studies have also shown the importance of prenatal care for preventing preterm birth. Some studies also identify the impact of socioeconomic variables, complications, and reproductive history on inadequate prenatal care and highlight the role of prenatal care as a determinant of maternal and perinatal indicators, citing such factors as low socioeconomic status, low schooling, and alcohol use in pregnancy [8,9,10,11], low maternal age, marital status, planning of the pregnancy, public healthcare services, high parity, previous premature childbirth [9,10,12], and unwanted pregnancy [13]. These determinants are part of the chain of factors associated with both inadequate prenatal care and negative outcomes such as preterm birth. We thus include the mediating role of prenatal care in the modeling.

Family vulnerability involves poverty, exploitation, abuse, and psychosocial and cultural factors. Another aspect in parallel is that families headed by women result partly from early or unwanted pregnancy, family instability, and abandonment [14]. The social vulnerability experienced by women may also be associated with a degree of emotional vulnerability involving feelings of abandonment, violence, and disempowerment which fails to ensure the conditions for adequate prenatal care and access to medicines [15]. Unplanned pregnancy is associated with smoking during pregnancy, using less folic acid, late initiation of prenatal care, and interruption of pregnancy [16].

Epidemiological studies on preterm birth routinely use hierarchical logistic regression models to identify risk factors for preterm birth that consider conceptual models to represent the relations between one variable and others in a causal model [17,18].

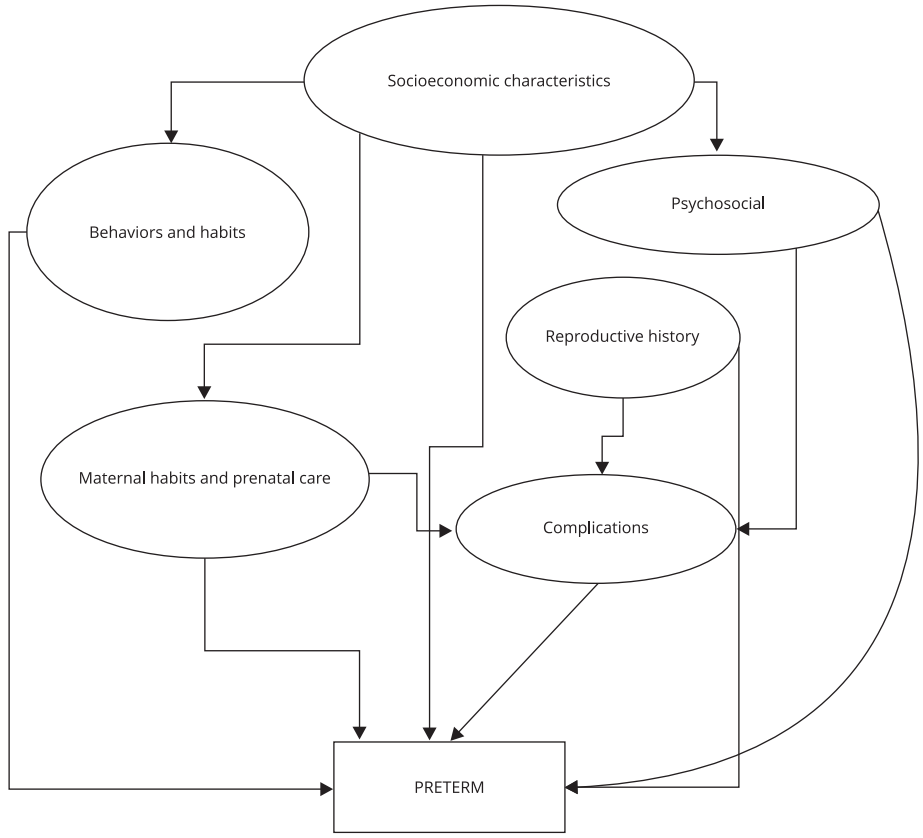
In such models, the variables are included simultaneously in the model and can have their effects overestimated or underestimated. When excluded, they may carry with them part of the information that makes the connection between the variables. Although hierarchical regression considers important elements such as temporal ordering and the logic of the relations between the variables, such models fail to adequately address the relations with confounding and mediating variables [19].

Structural equation models (SEM) are more appropriate, since they allow multiple simultaneous equations to incorporate confounding and mediation, besides incorporating latent variables for representing more complex measures that are not measurable with a single variable and that are created on the basis of covariance between observed variables. SEM minimize the effect of residual confounding in associations, especially in observational studies [20,21].

This study proposes to analyze the relations between variables and the outcome preterm birth via SEM, using a model tested with logistic regression and hierarchical selection of variables. We used data from a study in Londrina, Paraná State, Brazil, in 2009, headed by Silva et al. [2]. A model was redefined (Figure 1) based on the original study's theoretical and conceptual framework, and which incorporated latent variables. SEM also allowed including variables with a mediating effect on the exposure and outcome variables.

Figure 1

Theoretical and conceptual framework for factors associated with preterm birth.



Methods

The data are from a population-based case-control observational study of hospital births in Londrina (June 2006-March 2007) in which cases were defined as preterm births (< 37 gestational weeks) and controls were non-preterm births (≥ 37 weeks). The original study by Silva et al. [2] used all the variables in categorical form. In the current study, we opted to use the variables in continuous form in order to minimize the fact that categorical variables fail to display approximately normal distribution and may entail problems of asymmetry that affect the standard errors, estimates of residual variance, and the chi-square statistic used in the fit indices in SEM. The records with complete continuous variables totaled 296 cases and 329 controls, representing 90% and 89% of the original database, respectively.

In order to detect possible biases due to the difference in the number of records in the original database, the test of independence was used, based on Pearson's chi-square [22], which showed that the distributions of the variables in the two databases are statistically similar, indicating absence of bias in selection of the events.

The theoretical framework proposed by Silva et al. [2] was organized in blocks, which were transcribed in the current study to represent possible constructs (Figure 1). Block 1 in the original model is represented by maternal and family socioeconomic characteristics. Block 2 was separated into pregestational characteristics and maternal reproductive history. Block 3 was separated into psychosocial

conditions, maternal habits including physical activity and work, and prenatal care. Block 4 included maternal complications during the index pregnancy and block 5 covered fetal characteristics.

Based on this model, the following steps were adopted in the modeling:
(1) Measurement model: verification of the factors' composition via confirmatory factor analysis (CFA). Creation of latent variables: based on the constructs elaborated and suggested by the blocks of variables and the literature that defined the theoretical model, via CFA, considering the number of factors with eigenvalues greater than 1 [23].
(2) Structural model: relates latent variables, observed in the outcome preterm birth, drawing on models in which this outcome is binary. Direct and indirect effects on the outcome were obtained. Modification indices were used to identify relations not considered previously and improve the model's fit [24].
The weighted least square mean and variance adjusted estimator (WLSMV) was used with geomin rotation [25]. Having obtained the final model and tested the model's fits, it was re-estimated by robust maximum likelihood (MLR) in order to obtain odds ratio (OR) estimates. The fit indices used for the model were the Tucker-Lewis index (TLI; reference for good fit: TLI > 0.95), comparative fit index (CFI; reference: > 0.95), root mean square error of approximation (RMSEA; reference: < 0.05), and weighted root mean square residual (WRMR; reference: < 0.90). Significance was set at 5%. The analyses used the MPlus software package (https://www.statmodel.com/) and the graphs used CmapTools (https://cmap.ihmc.us/).

Maternal and family socioeconomic characteristics

These characteristics included the following variables: maternal age in years (< 19; 20 to 34; ≥ 35), per capita family income in minimum wage, number of household residents per room, mother's schooling in years, head-of-household's schooling in years, head-of-household's occupation (skilled; semi-skilled/manual; housekeeper; student/retired; unemployed), head-of-household's type of employment (formal; informal; not recorded), household location (non-slum; slum), type of family (nuclear – consisting of husband, wife, and children; only mother and children; non-nuclear; other family arrangements), mother living with husband/partner for less than two years (yes; no), person responsible for supporting household (mother; father; other), presence of elderly over 60 years of age (yes; no), mother's race/color (white/Asian/indigenous; black/brown), type of housing construction (finished; unfinished), maternal migration (yes; no), number of children under 10 years of age (yes; no), and head-of-household's age.

Psychosocial aspects

The gestational conditions covered psychosocial questions at the beginning and during pregnancy: attempted abortion (yes; no), planned pregnancy (yes; no), reactions to the pregnancy by the mother, father, and family and worries during the pregnancy (negative; positive), arguments/fights with the husband/partner (yes; no), separation (yes; no), and worries during the pregnancy (yes; no).

Maternal complications during the index pregnancy

Complications were: vaginal bleeding, hypertension, eclampsia, altered amniotic fluid volume, hospitalization during the pregnancy, anemia, diabetes, genital tract infection, vaginal discharge, placenta previa, and urinary tract infection (yes; no).

Maternal reproductive history

The variables used to measure maternal reproductive characteristics were parity, previous cesareans, previous preterm infant, previous low birth weight infant, maternal age, and assisted reproduction in this pregnancy (yes; no).

Work and physical activity

The variables used to measure activities were whether the mother worked during the pregnancy, whether the work required physical effort, and strenuous housework.

Prenatal care

Inadequacy of prenatal care involved the following observed variables: first consultation (1st trimester), number of consultations (minimum of 3), laboratory tests (urine, blood, ultrasound), procedures, and basic prenatal orientation. The categories were: adequate prenatal care (all items positive); inadequate level I (one or more negative answers); inadequate level II (three or more negative answers); or no prenatal care. The modeling considered two indicative variables to represent inadequacy of prenatal care: inadequacy of prenatal care I: if the prenatal care was inadequate (scored 1) and if adequate (scored 0); and inadequacy of prenatal care II: if the prenatal care was inadequate type II, which included mothers who had no prenatal care at all (scored 1) and if inadequate type I or adequate (scored 0).

Other variables

In order to cover all the variables selected by the original study [2], the following were included as observed variables: body mass index (BMI), physical activity, alcohol consumption, and multiple pregnancy.

Ethical aspects

The study complied with the ethical principles in *Resolution n. 196/1996* of the Brazilian National Health Council, which regulates research involving human subjects, and was approved by the Institutional Review Board of the Public Health Faculty/São Paulo University (protocol n. 404211/2013).

Results

Durante the modeling, changes were made to the initial model. Based on the blocks *Maternal and Family Socioeconomic Characteristics* and *Psychosocial Aspects*, the following latent variables were generated.

Maternal and family socioeconomic characteristics

Two factors were selected whose percentage of explained variance totaled 71.5%, with adequate fits in the model (RMSEA = 0.053; CFI = 0.988; TLI = 0.977) (Table 1). Variables whose correlation with the factors was less than 0.30 were excluded: mother's race/color, type of finishing on the housing, maternal migration, number of children under 10 years in the household, maternal age, head-of-household's age, and head-of-household's type of occupation. The two factors mentioned above represented socioeconomic vulnerability (SEV) and family vulnerability (FV), respectively. Based on the way the variables were categorized, higher scores for the latent variable SEV corresponded to lower maternal and head-of-household's schooling and higher rates of slum residence, since per capita family income has a negative sign, indicating that lower-income households have greater vulnerability. Meanwhile, higher scores for family vulnerability correspond to recent relations with the husband/partner, non-nuclear families, and presence of elderly in the household.

Table 1

Result of factor analysis with socioeconomic characteristics. Londrina, Paraná State, Brazil, 2009.

Variables	Factor 1	Factor 2
Low head-of-household's schooling (≤ 8 years)	**0.987**	-0.019
Low maternal schooling (≤ 8 years)	**0.795**	-0.022
Per capita income (minimum wage)	**-0.790**	-0.088
Number of persons per room	**0.735**	0.056
Type of dwelling (slum)	**0.648**	0.011
Non-nuclear family	0.158	**0.930**
Household headed by the mother	-0.069	**0.812**
Presence of elderly	0.005	**0.792**
Mother lives with husband/partner less than 2 years	-0.103	**0.692**

Psychosocial aspects

This latent variable aims to measure maternal psychological conditions affected by moments of worry and stress during pregnancy (Table 2). The variables attempted abortion, fights with husband/partners, separation, and worries during pregnancy were excluded due to low correlation (< 0.30) with the other variables. The selected variables were: planned pregnancy (yes; no) and reactions to the pregnancy by the mother, father, and family. A single factor was considered whose explained variance totaled 75.3%, with good fit indices (RMSEA = 0.056; CFI = 0.991; TLI = 0.972).

Based on the variables' categorization, higher scores on the unwanted pregnancy (UP) latent variable correspond to negative reactions to the pregnancy by the mother, father, and family. This factor also indicates emotional vulnerability, as measured by the level of support received by the mother after becoming pregnant.

Maternal reproductive history

This block was represented by the combination of parity, previous low birth weight infant, previous preterm infant, and maternal age associated with the outcome gradually with its OR and represents the interaction with these variables. (Table 3).

Maternal complications in the index pregnancy

In the original study's logistic regression model, the block related to maternal complications in the pregnancy showed an important fit in the multiple regression ^{2}, suggesting that more than one of these conditions could be present, and that some of them may have led to the mother's hospitalization during pregnancy. Thus, the variable to represent was the number of complications (0; 1 or more) (Table 3). Hospitalization also increased as the number of complications increased: 40% of the mothers with at least one complication had been hospitalized.

Work and physical activity: physical effort

Physical effort consisted of the following variables: whether the mother had worked during pregnancy, if the work required physical effort, and whether she had done strenuous housework (Table 3).

Table 2

Rotated factor loads (geomin method) – unwanted pregnancy. Londrina, Paraná State, Brazil, 2009.

Variables	Factor 1
Planned pregnancy	**0.911**
Mother's negative reaction or worries with the pregnancy	**0.969**
Father's negative reaction or worries with the pregnancy	**0.774**
Family's negative reaction or worries with the pregnancy	**0628**

Table 3

Distribution of cases and controls according to selected variables. Londrina, Paraná State, Brazil, 2009.

Variables	Cases		Controls		Total	
	n	%	n	%	n	%
Maternal reproductive history						
a) One or more pregnancies, no previous premature birth, and age 20 to 34 years	86	29.1	129	39.2	215	34.4
b) Nulliparous, age 13 to 19 years	35	11.8	44	13.4	79	12.6
c) One or more pregnancies, no previous premature birth, and age 35 years or older	22	7.4	38	11.6	60	9.6
d) Nulliparous, age 20 to 34 years	76	25.7	83	25.2	159	25.4
e) One or more pregnancies, no previous premature birth, no previous low birth weight, age 13 to 19 years	11	3.7	10	3.0	21	3.4
f) Nulliparous, age 35 years or older	9	3.0	2	0.6	11	1.8
g) One or more pregnancies, previous premature birth, previous low birth weight	57	19.3	23	7.0	80	12.8
Mother's behavioral activities: physical effort						
a) Mother does not work away from home, no strenuous housework	77	26.0	77	23.4	154	24.6
b) Mother does not work away from home, does strenuous housework	55	18.6	57	17.3	112	17.9
c) Mother works away from home without physical effort, no strenuous housework	94	31.8	92	28.0	186	29.8
d) Mother works away from home with physical effort, no strenuous housework	15	5.1	22	6.7	37	5.9
e) Mother works away from home without physical effort, does strenuous housework	28	9.5	47	14.3	75	12.0
f) Mother works away from home with physical effort, does strenuous housework	27	9.1	34	10.3	61	9.8

Final model

The complete model was developed on the basis of latent and observed variables (Table 4).

Table 4 presents the standardized coefficients and OR for the structural model's relations. The model's fit was satisfactory (CFI = 0.962; TLI = 0.952; WRMR = 1.005; RMSEA (95%CI) = 0.032 (0.026-0.039)).

The graphic representation of the final model (Figure 2) shows the standardized estimates. For some relations, the standardized estimates and OR are presented. For some variables, both values indicate the effects of inadequate prenatal care I and II.

The strongest direct effects on preterm birth were: complications (0.584), followed by inadequate prenatal care level II (0.555) and inadequate prenatal care level I (0.226), multiple pregnancy (0.214), reproductive history (0.091), BMI (0.146), alcohol consumption (0.114), walks (-0.133), and physical effort (-0.034). In addition to a direct effect (0.214), multiple pregnancy also showed an indirect effect on the outcome, via complications (0.224). When considering the total effect of multiple pregnancy (direct effect and via complications), the effect was stronger (0.382).

BMI during pregnancy also showed an indirect effect via complications (0.099), and its total effect on prematurity was 0.205. Besides a direct effect, alcohol consumption also showed an indirect effect

Table 4

Models' estimated parameters for preterm birth. Results with weighted least square mean and variance adjusted estimator (WLSMV) and robust maximum likelihood (MLR). Londrina, Paraná State, Brazil, 2009.

Variables and factors' indicators	WLSMV				MLR			
	Standardized estimate	Estimate	SE	p-value	Estimate	SE	p-value	OR
Measurement model								
SEV								
Persons per room	0.652	1.000	0.000	-	1.000	0.000	-	-
Per capita income	-0.877	-2.483	0.134	0.000	-2.412	0.170	0.000	-
Mother's schooling	0.841	4.798	0.476	0.000	7.274	0.843	0.000	-
Father's schooling	0.741	3.403	0.286	0.000	5.872	0.744	0.000	-
Type of dwelling	0.661	2.716	0.427	0.000	5.875	0.924	0.000	-
FV								
Living with husband/partner less than 2 years	0.669	1.000	0.000	-	1.000	0.000	-	-
Non-nuclear family	0.994	9.722	30.591	0.751	2.877	1.204	0.017	-
Presence of elderly	0.768	1.332	0.256	0.000	1.812	0.414	0.000	-
Relationship mother/guardian	0.776	1.367	0.218	0.000	1.503	0.245	0.000	-
UP								
Mother's reaction when learned of pregnancy	0.785	1.000	0.000	-	1.000	0.000	-	-
Father's reaction when learned of pregnancy	0.825	1.150	0.302	0.000	1.239	0.278	0.000	-
Family's reaction when learned of pregnancy	0.825	1.150	0.272	0.000	1.003	0.253	0.000	-
Structural model								
UP =								
FV	0.365	0.515	0.129	0.000	0.561	0.133	0.000	-
SEV	0.267	1.043	0.241	0.000	2.225	0.524	0.000	-
Mother's reproductive history	0.025	0.016	0.032	0.616	0.055	0.058	0.343	-
Alcohol consumption in pregnancy	0.166	0.983	0.368	0.007	1.265	0.590	0.032	-
Complications =								
Mother's reproductive history l	0.039	0.021	0.028	0.454	0.052	0.042	0.215	1.053
Physical effort	0.067	0.044	0.034	0.203	0.046	0.053	0.380	1.047
Walks	-0.020	-0.047	0.133	0.726	-0.267	0.199	0.179	0.765
BMI	0.099	0.155	0.076	0.041	0.295	0.127	0.021	1.343
Alcohol consumption in pregnancy	0.041	0.198	0.222	0.372	0.439	0.411	0.285	1.552
Multiple pregnancy	0.224	0.972	0.228	0.000	1.621	0.392	0.000	5.057
Worries	0.107	0.253	0.117	0.031	0.426	0.191	0.026	1.531
Inadequate prenatal care l =								
SEV	0.091	0.323	0.219	0.140	0.713	0.384	0.063	2.041
FV	0.088	0.112	0.099	0.256	0.163	0.116	0.160	1.177
UP	0.347	0.314	0.094	0.001	0.199	0.093	0.032	1.221
Mother's reproductive history	0.149	0.086	0.030	0.005	0.100	0.046	0.028	1.106
Physical effort	-0.092	-0.066	0.038	0.082	-0.050	0.060	0.399	0.951

(continues)

via UP and inadequate prenatal care, and its total effect was 0.126. Experiencing worries during pregnancy only showed an indirect effect (0.107) on prematurity via complications.

The latent variables UP and SEV showed direct effects on inadequate prenatal care. The latent variables showed indirect effects on preterm birth as follows: UP (0.228), SEV (0.139), and FV (0.123).

Table 4 (continued)

Variables and factors' indicators	WLSMV				MLR			
	Standardized estimate	Estimate	SE	p-value	Estimate	SE	p-value	OR
Structural model								
Inadequate prenatal care II =								
SEV	0.096	0.367	0.395	0.353	1.258	0.878	0.152	3.519
FV	-0.006	-0.008	0.175	0.963	0.105	0.200	0.601	1.110
UP	0.467	0.457	0.169	0.007	0.326	0.174	0.061	1.385
Mother's reproductive history	0.188	0.118	0.046	0.001	0.127	0.083	0.127	1.158
Physical effort	-0.159	-0.124	0.059	0.034	-0.154	0.130	0.237	0.857
Gestational age =								
SEV	-0.083	-0.551	0.611	0.367	0.007	0.383	0.986	1.007
FV	0.073	0.175	0.233	0.452	0.022	0.102	0.827	1.023
UP	-0.420	-0.714	0.462	0.122	-0.045	0.076	0.556	0.956
Mother's reproductive history	0.091	0.098	0.072	0.172	0.213	0.047	0.000	1.237
Inadequate prenatal care I	0.226	0.426	0.227	0.051	0.457	0.224	0.041	1.579
Inadequate prenatal care II	0.555	0.965	0.539	0.044	1.457	0.500	0.004	4.293
Complications	0.584	1.209	0.489	0.013	1.825	0.204	0.000	6.201
Multiple pregnancy	0.214	1.924	0.739	0.009	2.876	0.773	0.000	17.747
Physical effort	-0.034	-0.046	0.090	0.612	-0.146	0.064	0.023	0.864
Worries	0.040	0.198	0.227	0.384	0.424	0.218	0.052	1.527
Walks	-0.133	-0.641	0.300	0.032	-0.606	0.235	0.010	0.545
BMI	0.146	0.473	0.188	0.012	0.449	0.157	0.004	1.567
Alcohol consumption in pregnancy	0.114	1.147	0.555	0.039	1.131	0.393	0.004	3.099
Correlations								
FV with mother's reproductive history	0.065	0.116	0.096	0.225	0.226	0.148	0.127	-
SEV with mother's reproductive history	0.170	0.109	0.022	0.000	0.100	0.030	0.001	-
SEV with physical effort	-0.181	-0.093	0.022	0.000	-0.096	0.021	0.000	-
SEV with FV	0.175	0.051	0.016	0.001	0.079	0.027	0.003	-
SEV with BMI	0.147	0.032	0.009	0.001	0.025	0.010	0.014	-
SEV with walks	-0.228	-0.033	0.015	0.033	-0.033	0.007	0.000	-
BMI with walks	-0.043	-0.013	0.013	0.345	-0.011	0.011	0.024	-
Intercepts								
Persons per room	1.456	0.724	0.154	0.000	0.852	0.020	0.000	-
Per capita income	0.093	0.085	0.217	0.693	0.037	0.037	0.931	-

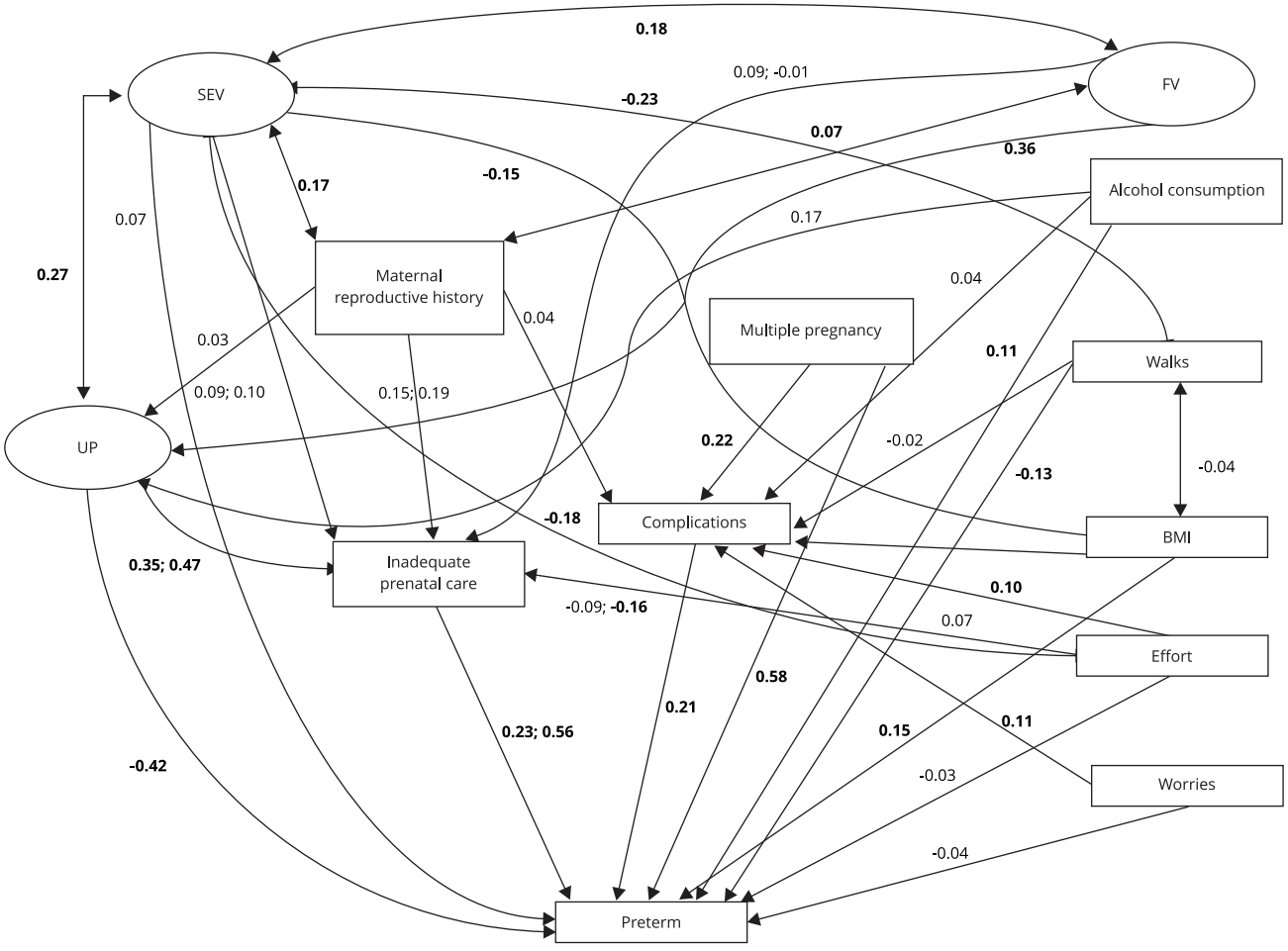
BMI: body mass index; FV: family vulnerability; OR: odds ratio; SE: standard error; SEV: socioeconomic vulnerability; UP: unwanted pregnancy.

Discussion

Prenatal care in the current model showed a direct effect with a positive sign on preterm birth, corroborating other studies [5,6]. In prenatal care, both the pregnant woman and her unborn child benefit from preventive follow-up, orientation, clarifications, and diagnosis of any altered health condition in the mother or fetus [26]. The effects of SEV, FV, reproductive history, and UP on prematurity occur indirectly through inadequate prenatal care, thus indicating the role of this variable's mediation on prematurity. Prenatal care plays an important role in negative pregnancy outcomes and reflects the mother's social, economic, and psychological conditions. According to a comprehensive analysis of quality indicators for prenatal care in Brazil, 15% of pregnant women receive adequate prenatal care and 60% receive all the recommended orientation and complementary tests [15]. Their living conditions and emotional and affective vulnerabilities can influence their knowledge on health during pregnancy [8,13,27,28]. Inadequate prenatal care was observed in young women with multiple previous

Figure 2

Resulting model for preterm birth.



BMI: body mass index; FV: family vulnerability; SEV: socioeconomic vulnerability; UP: unwanted pregnancy.

pregnancies, without husbands or partners, without paid employment, with little schooling, and with low socioeconomic status [29]. Based on our results, the relationship between these factors suggests the influence of factors of socioeconomic, family, and psychological vulnerability and maternal conditions on the adequacy of prenatal care.

Gestational complications, namely bleeding, high blood pressure, eclampsia, altered amniotic fluid volume, genital tract infection, and diabetes have a direct effect on preterm birth [1]. In addition to the direct effect, multiple pregnancy also displays an indirect effect via complications, which may be related to the occurrence of premature rupture of membranes or other maternal-fetal complications [1,2,5]. The variable maternal reproductive history showed a direct effect on the outcome. In the current study, this variable consisted of maternal age and reproductive history with the number of pregnancies and characteristics of any previous infants (low birth weight and prematurity), namely described factors [1,5,6,7].

The separation of socioeconomic and psychosocial factors in three factors aided the understanding of the various roles in this complex dimension. As a complement to the psychosocial characteristics, the variable worries expresses the nature of the stress experienced during pregnancy, which

was statistically significant in the original study. The current study did not identify a direct association with gestational age, but an indirect effect via complications, which points to another order of concerns, namely with health [1,2]. SEV, as a distal latent variable, showed correlations with both distal and proximal variables, such as physical effort, BMI, walks, maternal reproductive history, and FV. SEV was related indirectly to prematurity in the current study by two paths, the first via inadequate prenatal care and the second via UP. The socioeconomic dimension is quite complex and is not limited to unmet material needs of well-being but also to the denial of opportunities in social relations such as access to work and healthcare [30,31]. Both inadequate prenatal care and UP capture these vulnerabilities.

As for physical effort, the results here point to a negative association with prematurity, unlike other studies, including a systematic review in which either no effect was observed (in the majority of studies) or a moderate effect from work conditions involving physical effort was seen. The variable physical effort, as constructed here, can translate the mother's life phase in relation to her full productive activity/work, whether due to her age or household per capita income, pointing not to SEV (correlation with the variable SEV verified in the model). Meanwhile, Behrman & Butler [28] found that strenuous work can also be seen as an indicator of favorable socioeconomic circumstances, i.e., the capacity to have and keep a job, along with the employment benefits, while also signaling the psychological satisfaction resulting from some types of work. Thus, this result in the current study is not fully explained, and other approaches are needed to better understand the role of physical effort. The same pattern was also seen in the variable taking walks or other forms of physical activity, that is, a direct and negative effect on prematurity.

The BMI variable had a direct effect on gestational age at birth and an indirect effect via complications, corroborating the findings by Padilha et al. [32]. The current study also found a correlation between BMI and SEV, indicating BMI levels outside the normal range, namely < 19kg/m^2 (underweight) and ≥ 30kg/m^2 (obese) in mothers with greater SEV.

This same pattern was also seen in alcohol consumption during pregnancy, with direct and indirect effects on gestational age at birth. The indirect effect was via UP and inadequate prenatal care. Other studies have shown an association between heavy alcohol consumption and prematurity, but the mechanisms were not clear [1,28,33], while still others found no association [34].

The current model also indicated a path to prematurity via UP and inadequate prenatal care, and no references were found that discuss these possible relations. Future studies could explore these paths better. Again, the role of prenatal care as a channel of care for the mother and infant deserves attention.

The original case-control study [2] used a theoretical framework designed in hierarchical blocks using a logistic regression model [35] in which blocks of variables were grouped according to some common characteristic. The current study was based on this model by considering the structure created by including non-observed (latent) variables. In principle, it is possible to create a structural equation model drawing on a hierarchical model, since the structure in blocks bears an important element of the relations between these variables and the outcome: temporality, characterized by the concepts "distal", "intermediate", and "proximal".

However, there were difficulties in rewriting the model: the blocks combine variables that are not always correlated, so that each relationship should be reassessed for inclusion in the model.

In addition to the low correlations between the variables comprising the factors, various other problems may arise with the use of the structural equation model: estimation problems, non- convergence, lack of identification, non-positive matrix, and estimation of negative variances are not uncommon [26].

For some variables, estimation problems occurred due to empty boxes (zero) in the bivariate combinations, such as "previous pregnancy" and "nulliparous mothers", in which structurally there is a zero in this intersection, which prevents estimation of the covariance and requires using the solution proposed by Muthén & Muthén [36], namely to combine a new variable, that is, to create an interaction variable. Its relationship to the outcome was taken into account (via OR). An alternative is to use a perturbation in the box, i.e., the zero is replaced by a small value [37].

Estimation of the OR allows the known interpretation in epidemiology. Obtaining the OR model after having generated the model with WLSMV was the alternative found to perform the most

adequate estimation of the data [38]. Muthén et al. [39] reported that WLSMV was developed with small and moderate samples. Beauducel & Herzberg [40] also reported more satisfactory results with WLSMV in variables with few categories (2 or 3, compared to 4 or more).

The study's limitations include its observational design, which does not allow conclusions as to causality. The number of individuals was small given the number of relations studied, although this was not a problem for obtaining the models' fits. And the study was not designed to observe the relations in the form proposed by the modeling.

From the methodological point of view, the SEM was built based on a model that had already been built (the original model), drawing on hierarchical regression, presenting advantages to the extent that it can be used as an initial model and allowed including paths that represent the relations between the model's variables (between them and the outcome), besides identifying proposed mediations. The proposed alterations to the initial model aimed to grasp the variables involved in the original study. The process of building the SEM highlighted: (1) the need to reanalyze the role of the variables when choosing to use the SEM; (2) that complex models require studies that already consider the creation of latent variables in their formulation; (3) that the combination of variables to form a new variable is a resource in the absence of well-defined prior latent variables and that solves problems of empty boxes when cross-analyzing variables; and (4) that the same variables were significant using the different estimators (WLSMV and MLR) in the model's structural component.

Conclusion

It was possible to use the result of the original work to develop the SEM based on the revision of the theoretical model, knowing the relations with the outcome and the logistic regression model. Application of the proposed methodology identified the presence of constructs (SEV, UP, FV), verified inadequacy of prenatal care as a mediator, and identified direct and indirect effects of variables on the outcome preterm birth.

Contributors

A. A. Oliveira participated in the project's conception, data analysis and interpretation, and writing of the article. M. F. Almeida participated in the data analysis and interpretation and writing of the article. Z. P. Silva and P. L. Assunção participated in the analysis and revision of the text. A. M. R. Silva and H. G. Santos participated in the revision of the text. G. P. Alencar participated in the project's conception and in the data analysis and interpretation, conducted a critical revision of the content, and approved the final version for publication.

Additional informations

ORCID: Adelaide Alves de Oliveira (0000-0002-8997-8886); Marcia Furquim de Almeida (0000-0003-0052-1888); Zilda Pereira da Silva (0000-0003-4648-113X); Paula Lisiane de Assunção (0000-0001-6261-1657); Ana Maria Rigo Silva (0000-0002-9527-4914); Hellen Geremias dos Santos (0000-0002-6446-8660); Gizelton Pereira Alencar (0000-0002-2354-9050).

Acknowledgments

The authors wish to thank the São Paulo State Research Foundation (Fapesp; grant n. 2010/20784-8) for the research funding.

References

1. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet 2008; 371:75-84.
2. Silva AMR, Almeida MF, Matsuo T, Soares DA. Fatores de risco para nascimentos pré-termo em Londrina, Paraná, Brasil. Cad Saúde Pública 2009; 25:2125-38.
3. Hair Jr. JF, Anderson RE, Tatham RL, Black WC. Análise multivariada de dados. 6th Ed. Porto Alegre: Bookman; 2009.
4. Assunção PL, Novaes HMD, Alencar GP, Melo ASO, Almeida MF. Fatores associados ao nascimento pré-termo em Campina Grande, Paraíba, Brasil: um estudo caso-controle. Cad Saúde Pública 2012; 28:1078-90.
5. Silveira MF, Victora CG, Barros AJD, Santos IS, Matijasevich A, Barros FC. Determinants of preterm birth: Pelotas, Rio Grande do Sul State, Brazil, 2004 birth cohort. Cad Saúde Pública 2010; 26:185-94.
6. Aragão VMF, Silva AAM, Aragão LF, Barbieri MA, Bettiol H, Coimbra LC, et al. Risk factors for preterm births in São Luís, Maranhão, Brasil. Cad Saúde Pública 2004; 20:57-63.
7. Bezerra LC, Oliveira SMJV, Latorre MRDO. Prevalência e fatores associados à prematuridade entre gestantes submetidas à inibição de trabalho de parto prematuro. Rev Bras Saúde Matern Infant 2006; 6:223-9.
8. Goudard MJF, Simões VMF, Batista RFL, Queiroz RCS, Alves MTSSB, Coimbra LC, et al. Inadequação do conteúdo da assistência pré-natal e fatores associados em uma coorte no nordeste brasileiro. Ciênc Saúde Coletiva 2016; 21:1227-38.
9. Feijen-de Jong EI, Jansen DEMC, Baarveld F, van der Schans CP, Schellevis FG, Reijneveld SA. Determinants of late and/or inadequate use of prenatal healthcare in high-income countries: a systematic review. Eur J Public Health 2011; 22:904-13.
10. Carvalho VCP, Araújo TVB. Adequação da assistência pré-natal em gestantes atendidas em dois hospitais de referência para gravidez de alto risco do Sistema Único de Saúde, na cidade de Recife, Estado de Pernambuco. Rev Bras Saúde Matern Infant 2007; 7:309-17.
11. Ribeiro ERO, Guimarães AMDN, Bettiol H, Lima DDF, Almeida MLD, Souza L, et al. Risk factors for inadequate prenatal care use in the metropolitan area of Aracaju, Northeast Brazil. BMC Pregnancy Childbirth 2009; 9:31.
12. Trevisan MR, Lorenzi DS, Araújo NM, Ésber K. Perfil da assistência pré-natal entre usuárias do Sistema Único de Saúde em Caxias do Sul. Rev Bras Ginecol Obstet 2002; 24:293-9.
13. Kramer MS, Demissie K, Yang H, Platt RW, Sauvé R, Liston R. The contribution of mild and moderate preterm birth to infant mortality. Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System. JAMA 2000; 284:843-9.
14. Koga D. Medida de cidades: entre territórios de vida e territórios vividos. São Paulo: Cortez Editora; 2003.
15. Tomasi E, Fernandes PAA, Fischer T, Siqueira FCV, Silveira DS, Thumé E, et al. Qualidade da atenção pré-natal na rede básica de saúde do Brasil: indicadores e desigualdades sociais. Cad Saúde Pública 2017; 33:e00195815.
16. Borges ALV, Cavalhieri FB, Hoga LAK, Fujimori E, Barbosa LR. Planejamento da gravidez: prevalência e aspectos associados. Rev Esc Enferm USP 2011; 45:1679-84.
17. Rothman KJ, Greenland S. Modern epidemiology. 3rd Ed. Philadelphia: Lippincott-Raven; 2012.
18. Susser M. Causal thinking in the health sciences concepts and strategies of epidemiology. New York: Oxford University Press; 1973.
19. Silva AAM, Vasconcelos AGG, Bettiol H, Barbieri MA. Socioeconomic status, birth weight, maternal smoking during pregnancy and adiposity in early adult life: an analysis using structural equation modeling. Cad Saúde Pública 2010; 26:15-29.
20. Amorim LDAF, Fiaccone RL, Santos CAST, Santos TN, Morais LTLP, Oliveira NS, et al. Modelagem de equações estruturais em epidemiologia. Cad Saúde Pública 2010; 26:2251-62.
21. Hoyle RH. Structural equation modeling: concepts, issues, and applications. Thousand Oaks: SAGE; 1995.
22. Bussab WO, Morettin PA. Estatística básica. 8ª Ed. São Paulo: Saraiva; 2013.
23. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6th Ed. New Jersey: Prentice Hall; 2007.
24. Schumacker RE, Lomax RG. A beginner's guide to structural equation modeling. 3rd Ed. New York: Routledge; 2010.
25. Byrne BM. Structural equation modeling with Mplus: basic concepts, applications, and programming. New York: Routledge; 2011.
26. Rios CTF, Vieira NFC. Ações educativas no pré-natal: reflexão sobre a consulta de enfermagem como um espaço para educação em saúde. Ciênc Saúde Coletiva 2007; 12:477-86.
27. Líbera BD, Saunders C, Santos MMAS, Rimes KA, Brito FRSS, Balão MR. Avaliação da assistência pré-natal na perspectiva de puérperas e profissionais de saúde. Ciênc Saúde Coletiva 2011; 16:4855-64.
28. Behrman RE, Butler AS. Preterm birth: causes, consequences and prevention. Washington DC: National Academies Press; 2007.
29. Domingues RMSM, Viellas EF, Dias MAB, Torres JA, Theme-Filha MM, Gama SGN, et al. Adequação da assistência pré-natal segundo as características maternas no Brasil. Rev Panam Salud Pública 2015; 37:140-7.

30. Lima AMC. Modelagem de equações estruturais: uma contribuição metodológica para o estudo da pobreza [Dissertação de Mestrado]. Salvador: Universidade Federal da Bahia; 2005.
31. Instituto de Pesquisa Econômica Aplicada. Pobreza multidimensional no Brasil. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada; 2006.
32. Padilha PC, Saunders C, Machado RCM, Silva CL, Bull A, Sally EOF, et al. Associação entre o estado nutricional pré-gestacional e a predição do risco de intercorrências gestacionais. Rev Bras Ginecol Obstet 2007; 29:511-8.
33. Albertsen K, Andersen AMN, Olsen J, Gronbaek M. Alcohol consumption during pregnancy and the risk of preterm delivery. Am J Epidemiol 2004; 159:155-61.
34. Peacock JL, Bland JM, Anderson HR. Preterm delivery: effects of socioeconomic factors, psychological stress, smoking, alcohol, and caffeine. BMJ 1995; 311:531-5.
35. Victora CG, Huntly SR, Fuchs SC, Olinto MTA. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. Int J Epidemiol 1997; 26:224-7.
36. Muthén LK, Muthén BO. Mplus user's guide. 6th Ed. Los Angeles: Muthén & Muthén; 2010.
37. Oberski DL. lavaan.survey: an R package for complex survey analysis of structural equation models. J Stat Softw 2014; 57:1-27.
38. Brown TA. Confirmatory factor analysis for applied research. New York: The Guilford Press; 2015.
39. Muthén BO, du Toit SHC, Spisic D. Robust inference using weighted least squares and quadratic estimating equations in latent variable modelling with categorical and continuous outcomes. Los Angeles: University of California, Los Angeles; 1997. (Technical Report).
40. Beauducel A, Herzberg PY. On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. Struct Equ Modeling 2006; 13:186-203.

Resumo

O estudo propôs a aplicação da modelagem com equações estruturais (MEE) para estudar variáveis associadas ao nascimento pré-termo com base em um modelo teórico analisado previamente pela regressão logística hierarquizada. Os dados foram provenientes da pesquisa observacional do tipo caso-controle populacional sobre nascidos vivos hospitalares de mães residentes em Londrina, Paraná, Brasil (junho de 2006 a março de 2007). Para a MEE foi considerada a associação de características socioeconômicas e aspectos psicossociais sobre história reprodutiva, trabalho e atividade física, intercorrências durante a gestação e características fetais. Considerou-se, ainda, a relação dessas associações sobre o desfecho nascimento pré-termo mediado pela adequação da assistência pré-natal. Foram utilizados estimadores de mínimos quadrados ajustados pela média e variância (WLSMV), para dados categóricos, e a máxima verossimilhança robusta (MLR), para obter razões de chances. Foram criadas três variáveis latentes: vulnerabilidade socioeconômica, vulnerabilidade familiar e não aceitação da gravidez. O efeito da vulnerabilidade socioeconômica, da família e da não aceitação da gravidez sobre a prematuridade ocorreu de modo indireto por meio da inadequação da assistência pré-natal. A metodologia proposta possibilitou utilizar construtos, verificar o papel de mediação da inadequação da assistência pré-natal e identificar efeitos diretos e indiretos das variáveis sobre o desfecho nascimento pré-termo.

Modelos Estatísticos; Modelos Logísticos; Fatores de Risco; Nascimento Prematuro

Resumen

Este estudio propuso la aplicación de modelos de ecuaciones estructurales (SEM) para investigar las variables asociadas con el parto prematuro basándose en un modelo teórico previamente analizado mediante regresión logística jerárquica. Los datos provienen de un estudio observacional de casos y controles de base poblacional de nacidos vivos en hospitales de madres que residen en Londrina, estado de Paraná, Brasil (junio de 2006 a marzo de 2007). Para el SEM, el estudio consideró la asociación entre las características socioeconómicas y los aspectos psicosociales relacionados con el historial reproductivo, el trabajo y la actividad física, las complicaciones durante el embarazo y las características fetales. También consideró la relación entre estas asociaciones y el parto prematuro mediado por la adecuación de la atención prenatal. Se utilizó el estimador de los mínimos cuadrados ponderados ajustados por la media y variancia (WLSMV) para datos categóricos y la probabilidad máxima robusta (MLR) para los odds ratios. Se crearon tres variables latentes: vulnerabilidad socioeconómica, vulnerabilidad familiar y embarazo no deseado. El efecto de la vulnerabilidad socioeconómica y familiar y el embarazo no deseado en la prematuridad ocurrió indirectamente por la insuficiencia de la atención prenatal. La metodología propuesta permitió usar constructos, verificar el papel de la mediación por la insuficiencia de la atención prenatal e identificar los efectos directos e indirectos de las variables sobre el resultado "parto prematuro".

Modelos Estadísticos; Modelos Logísticos; Factores de Riesgo; Nacimiento Prematuro

Submitted on 05/Dec/2017
Final version resubmitted on 03/Jul/2018
Approved on 06/Jul/2018