

Influência dos fatores de situação socioeconômica, de aceitação da gravidez e da assistência pré-natal na mortalidade fetal:
análise com modelagem de equações estruturais

Gizelton Pereira Alencar

Tese apresentada na área de concentração de Epidemiologia da Faculdade de Saúde Pública da Universidade de São Paulo para obtenção do Grau de Doutor.

Área de concentração: Epidemiologia

Orientadora: Prof. Dr. Marcia Furquim de Almeida

São Paulo

2009

Autorizo, exclusivamente para fins acadêmicos e científicos, a reprodução total ou parcial desta tese, por processos fotocopiadores.

Assinatura:

Data:

AGRADECIMENTOS

Gostaria de agradecer a todos que, direta ou indiretamente, me ajudaram durante este período em que fiz o doutorado. Em especial, gostaria de agradecer:

À professora Marcia Furquim de Almeida, minha orientadora, pelos ensinamentos e oportunidades oferecidas, pela generosidade e coragem de ousar trabalhar com novos conceitos e idéias, e por sua amizade e compreensão nos momentos difíceis pelos quais passei.

Ao professor José Maria Pacheco de Souza pelas valiosas sugestões durante todo o meu doutorado e pelos ensinamentos sobre estatística e epidemiologia, e na vida acadêmica.

Ao professor Clovis de Araujo Peres pelas valiosas sugestões durante meu doutorado e pela grande experiência aprendida com ele na UNIFESP e em tantos outros momentos.

À professora Rita Barradas Barata, pelo interesse e disposição em discutir o meu trabalho, por oferecer caminhos para a reflexão e pelas sugestões dadas.

À professora Silvia Nagib Elian, minha professora desde a graduação, pelo interesse e disposição em discutir o meu trabalho e pelas sugestões propostas.

Às professoras Laura Rodrigues e Oona Campbell, da LSHTM, pela receptividade carinhosa e disposição em discutir o meu trabalho. *Thank you very much.*

Aos professores Maria Novaes, Ivan França e Arnaldo Siqueira, pelos ensinamentos nas nossas reuniões do grupo de pesquisa que influenciaram enormemente este trabalho.

Aos professores e funcionários do Departamento de Epidemiologia que me apoiaram no decorrer do tempo em que estive na FSP.

À Cidinha, Renilda, Vânia e Viviane pelo apoio e ao pessoal da Biblioteca da FSP.

À Célia, pelo carinho, pela compreensão, por estar ao meu lado e por todo o apoio e incentivo na confecção deste trabalho.

Aos meus pais José Airton e Gizelane, meus avós, sogros e cunhados, sobrinhos e demais familiares, pelo apoio.

À Lane Alencar, Rosana Alencar, Marcelo Rocha, Marcus Vinicius, Daniela Schoeps, Gustavo Pereira, José Ricardo, Rafael Cunha, Andréia Santos, Luiza Cristina, Mazé e Luzia, e tantos outros amigos da ETESP, IME, FSP e LSHTM, pelo apoio, amizade, ensinamentos e paciência.

Aos amigos da SVS, em Brasília, com quem trabalhei e aprendi muito sobre a saúde no Brasil.

À CAPES pelos 8 meses de bolsa-sanduíche que propiciou minha ida à LSHTM.

Resumo

Alencar GP. **Influência dos fatores de situação socioeconômica, de aceitação da gravidez e da assistência pré-natal na mortalidade fetal: análise com modelagem de equações estruturais.** São Paulo; 2009. [Tese de Doutorado - Faculdade de Saúde Pública da USP].

Introdução. As hipóteses levantadas sobre as relações entre as variáveis de exposição e o óbito fetal, via diversos mecanismos, são desenhadas em um diagrama a partir de um quadro conceitual feito pelo pesquisador. Testes sobre essas relações podem ser feitos por meio da modelagem por equações estruturais (MEE).

Objetivo. Este estudo objetiva compreender o papel da situação socioeconômica (SSE), da não-aceitação da gravidez (nAdG) e assistência pré-natal na mortalidade fetal, na região sul do município de São Paulo, em 2000.

Métodos. Foram construídas variáveis latentes para representar a situação socioeconômica e a não-aceitação da gravidez, a partir de dados da pesquisa do tipo caso-controle sobre mortalidade fetal no município de São Paulo, feita em 2000. Foi testado um modelo para o desfecho óbito fetal com as variáveis latentes, a assistência pré-natal, intercorrências (hipertensão, diabetes gestacional, sangramento vaginal) e retardo de crescimento intra-uterino (RCIU) por meio do MEE. A SSE foi formada por escolaridades da mãe e do chefe da família, classificação socioeconômica e renda per capita; a nAdG foi formada a partir das reações da mãe, pai e família, tentativa de aborto e se a gravidez foi planejada.

Resultados. O modelo final apontou a relação entre a situação socioeconômica e não-aceitação da gravidez sobre a mortalidade fetal, passando pelo cuidado de pré-natal inadequado, modelo no qual intercorrências e RCIU também tem efeito sobre o desfecho.

Conclusões. Os resultados sugerem que aceitação negativa da gravidez e o efeito da situação socioeconômica desfavorável, mediados pela atenção de pré-natal, podem ser amenizadas com uma maior qualidade nas visitas de pré-natal, especialmente em mães de maior vulnerabilidade social.

Abstract

Alencar GP. Influência dos fatores de situação socioeconômica, de aceitação da gravidez e da assistência pré-natal na mortalidade fetal: análise com modelagem de equações estruturais. [Influence of socioeconomic situation and acceptance of pregnancy factors and prenatal care on fetal mortality: analysis through structural equation modeling]. São Paulo (BR); 2009. [PhD Thesis - Faculdade de Saúde Pública da Universidade de São Paulo, Brazil].

Introduction. The hypothesis about the relations among exposure variables and the fetal death are written as a diagram designed from a conceptual framework given by the researcher. Tests on these relations can be performed using structural equation modeling (SEM).

Objective. This study aim to understand the role of socioeconomic situation (SES), non-acceptance of pregnancy (nAoP) and prenatal care on fetal mortality, in region south of the city of São Paulo, in 2000.

Methodology. The measurement models were created to the latent variables SES and nAoP, from case-control study on fetal mortality in Sao Paulo city, in 2000. It was tested a model with SES, nAoP, prenatal care, medical complications (hypertension, gestational diabetes, vaginal bleeding), intra-uterine growth retardation (IUGR) on the outcome, performed by SEM.

Results. The SES was composed by mothers' and household head's years of study, socioeconomic classification index, and income per capita income; The AoP was composed by reactions of mother, father and family, attempted abortion and planned gestation. The final model showed the effect of SES and AoP on fetal death mediated by prenatal care, in a model where medical complications and IUGR also has an effect.

Conclusion. The results suggests that negative acceptance of pregnancy and unfavorable socioeconomic situation, mediated by prenatal care, could be minimized improving the quality of prenatal care especially to those who are more social vulnerable.

Índice

1 INTRODUÇÃO	1
1.1 Mortalidade fetal	2
1.2 Multicausalidade	6
1.3 Variável de confusão e variável mediadora (interveniente).....	9
1.4 Modelos de Equações Estruturais (MEE)	12
1.4.1 Representação gráfica - diagrama de caminhos	13
1.4.2 A análise fatorial no contexto de MEE	15
1.4.3 Limitações da análise	16
1.4.4 O MEE e as variáveis categóricas ordinais e binárias – estimação.....	17
1.4.5 Etapas do MEE.....	20
1.4.6 Interpretação.....	21
2 OBJETIVOS	23
3 METODOLOGIA	24
3.1 Material	24
3.1.1 Fonte de dados	24
3.1.2 Tipo de estudo.....	24
3.1.3 Período de referência e população de estudo	24
3.1.4 Variáveis de estudo	25
3.2 Métodos.....	26
3.2.1 Passos adotados na modelagem	28
4 RESULTADOS.....	32
4.1 Modelagem de equações estruturais para a mortalidade fetal.....	32
4.2 Discussão	51
4.2.1 Discussão sobre o caminho causal para óbitos fetais.....	51
4.2.2 Discussão sobre a metodologia utilizada	55

4.3 Conclusões	58
5 REFERÊNCIAS	59
6 QUESTÕES ÉTICAS	68
ANEXO 1 – ESCRREVENDO O MODELO DE EQUAÇÕES ESTRUTURAIS	69
ANEXO 2 – IDENTIFICABILIDADE EM MODELOS DE EQUAÇÕES ESTRUTURAIS	76
ANEXO 3 – ESTIMAÇÃO PARA DADOS BINÁRIOS	78
ANEXO 4 – MEDIDAS DE AJUSTE.....	83
ANEXO 5 – TABULAÇÕES DAS VARIÁVEIS OBSERVADAS	85
ANEXO 6 – COMITÊ DE ÉTICA EM PESQUISA	87

1 INTRODUÇÃO

A concepção multicausal dos agravos em saúde é parte do desenvolvimento teórico e metodológico da Epidemiologia no século XX. Esta concepção foi, inicialmente, apresentada por MacMahon (1970), o que permitiu na década de 1970 uma melhor compreensão da etiologia das doenças não-transmissíveis e, desse modo, possibilitou o desenvolvimento de instrumentos para aprimorar sua prevenção e controle (Barata 2005).

A concepção de multicausalidade da doença, contudo, traz consigo desafios teóricos e metodológicos de como tratar as diversas exposições envolvidas em determinado desfecho. Na década de 1970, com a crescente incorporação das técnicas estatísticas de tratamento de dados e a possibilidade do uso de computadores para viabilizar as análises surgiram importantes marcos no conhecimento da Epidemiologia. Susser (1973) traz uma extensa discussão sobre a multicausalidade, considerando o desfecho como variável dependente e as exposições como variáveis independentes e introduz a noção de variável interveniente (que será discutida mais à frente), variável essa que se encontra associada à exposição e ao desfecho e se encontra no caminho causal da ocorrência deste evento.

O avanço da microcomputação e de aplicativos estatísticos com interfaces mais amigáveis possibilitou que as análises estatísticas ficassem mais acessíveis e ágeis; paralelamente, houve intenso desenvolvimento de estudos de base individual (caso-controle, coorte e ensaios clínicos) e o emprego de análises multivariadas. A modelagem estatística tornou-se cada vez mais freqüente e, associada a isso, há a incorporação de novas metodologias. O avanço teórico e metodológico possibilitou intensa discussão no trabalho com diferentes níveis de exposição e tem trazido contribuição importante para a compreensão dos fatores de risco de diversas doenças.

A operacionalização do conceito de causalidade é importante para a epidemiologia, pois a necessidade de identificação das causas de um fenômeno ou de seu mecanismo é uma das bases do avanço do conhecimento. A quantificação do risco foi a forma

utilizada para se estudar a causalidade (Susser 1973) e levou à discussão e o desenvolvimento de metodologias estatísticas.

Devido ao declínio da mortalidade infantil e perinatal e à importância crescente do componente fetal, há necessidade de elaboração de estudos que busquem conhecer as variáveis associadas ao óbito fetal de um ponto de vista que considere relações mais complexas entre essas variáveis.

1.1 Mortalidade fetal

O número de óbitos neonatais no mundo gira em torno de 4 milhões. Para cada recém-nascido que morre na primeira semana depois do nascimento, um outro bebê nasce morto (óbito fetal intraparto ou anteparto) (OMS 2006) e, assim, os óbitos fetais representam praticamente metade dos óbitos perinatais. Atualmente no mundo, a taxa de mortalidade fetal varia de 3 a 5 óbitos por 1000 nascimentos, nos Estados Unidos, chegando a 30 por 1000 nascidos vivos em países da África subsahariana. Taxas de mortalidade fetal entre 10 e 15 por 1000 nascidos vivos são encontrados em países de nível intermediário de desenvolvimento (OMS 2006). Em 2000, segundo as informações do estudo sobre mortalidade perinatal na região Sul do município de São Paulo, a estimativa da taxa de mortalidade fetal era de 8,4 óbitos fetais por 1000 nascidos vivos (Almeida et al 2007).

Dentre as variáveis associadas aos óbitos fetais anteparto tem-se: condições socioeconômicas, demográficas e psicossociais, tais como: tipo de ocupação do chefe de família, carga social e psicológica das mães, bem como mães sem união marital e de baixa escolaridade (Chalumeau 2002, Künzel 2003, Stephansson 2001). Algumas características maternas antes da gestação, tais como idade materna maior do que 35 anos, alta paridade, curtos intervalos inter-partais, altura da mãe abaixo de 1,5m e índice de massa corpórea ($IMC > 29,0 \text{ kg/m}^2$) (Chalumeau 2002, Conde-Agudelo 2000, Künzel 2003). Também encontram-se associadas ao óbito fetal condições desfavoráveis em gestações anteriores (Chalumeau 2002, Surkan 2004, Zhang 2004), a presença de infecções na gravidez (sífilis ou outras), e intercorrências tais como diabetes, hipertensão, pré-eclâmpsia, sangramento vaginal, problemas na placenta e

fumo (Chalumeau 2002, Künzel 2003). Cuidado pré-natal inadequado (Chalumeau 2002, Conde-Agudelo 2000) e algumas características fetais (RCIU, malformações congênicas) (Chalumeau 2002, Conde-Agudelo 2000, Zhang 2004) foram identificados como fatores de risco. Também estão associadas à ocorrência do óbito fetal a poluição e a alimentação deficiente (Pereira et al. 1998, OMS 2006).

Segundo McClure (2007), 97% dos óbitos fetais ocorrem em países em desenvolvimento. Os óbitos fetais têm sido pouco estudados e raramente têm sido considerados para melhorar as condições adversas na gravidez (OMS 2006). Neste estudo apresenta resultados para países com condições bastante desfavoráveis, como República do Congo, Guatemala, Índia, Zâmbia e Paquistão e, também, para a Argentina, país considerado pelo estudo como um país de condições intermediárias. Na Argentina, 99,9% dos partos analisados eram hospitalares e 68,9% foram realizados por um médico. Os fatores de risco associados aos óbitos fetais foram: idade materna acima de 35 anos no momento do parto, baixo nível de escolaridade, primíparas ou grande múltiparas (4 ou mais). A ausência de atenção pré-natal ou a baixa qualidade da atenção recebida no momento do parto e nascimentos não hospitalares também foram identificados como risco em todos os países estudados. Com relação às características do feto, verificou-se que o sexo masculino, nascimentos de pré-termo e de baixo peso ao nascer apresentaram maior risco para ocorrência do óbito fetal. Além disso, identificou-se que a prevalência de malformações congênicas era inferior a 1%, no momento do parto.

Kramer et al. (2001), em artigo sobre os possíveis fatores de risco para nascimentos de pré-termo, sugerem que a identificação do efeito de algumas variáveis dificilmente pode ser avaliada apenas por meio de técnicas de regressão e, dentre esses efeitos, menciona a aceitação da gravidez, a presença ou ausência de suporte social durante a gestação e a situação socioeconômica das mães e famílias. Misra et al. (2001) também sugerem que a presença de relações entre variáveis que foram denominadas de biomédicas, tais como presença de doenças crônicas pré-existentes, complicações agudas da gravidez como, por exemplo, pré-eclampsia, e comportamentos negativos de saúde, como hábito de fumar e assistência inadequada de pré-natal, e situação socioeconômica desfavorável (como privação material); e

condições psicossociais (presença de estresse, depressão e suporte social) afetam o desfecho de nascimento de pré-termo.

O estudo de fatores de risco para óbitos fetais anteparto na região sul de São Paulo (Almeida et al. 2007) identificou os seguintes fatores de risco: das características socioeconômicas e demográficas das mães e das famílias: mãe sem união marital e baixa escolaridade da mãe; característica da mãe antes da gravidez: nascimento de baixo peso em gestação anterior; das condições durante a gestação: sangramento vaginal, hipertensão gestacional e diabetes, além de pré-natal inadequado. Dentre as características do feto: malformações congênitas e retardo de crescimento intra-uterino foram identificados como fatores de risco para o óbito fetal.

No entanto, diversas questões foram levantadas na análise de dados e permaneceram sem uma resposta clara. Verificou-se que a assistência pré-natal inadequada era mais freqüente entre as mães de baixa escolaridade, o que poderia sugerir que a assistência pré-natal poderia ser uma variável mediadora, isto é, parte das condições de vida precárias das mães poderia se expressar por meio da atenção pré-natal inadequada. Os resultados mostraram também que a não aceitação da gravidez aumentava o risco para óbitos fetais, porém este efeito desaparecia ao se acrescentar as variáveis relativas às condições da gestação, em especial, ao se incluir a atenção pré-natal. Por outro lado, identificou-se existir uma associação entre a não realização ou inadequação pré-natal e a reação negativa da mãe em relação à gestação. Levanta-se, então, a seguinte questão: a não aceitação da gestação pode levar à realização de pré-natal inadequado e, como consequência, constituir-se em risco para óbitos fetais?

Para tentar responder essas questões, decidiu-se empregar a modelagem de equações estruturais (MEE). Essa metodologia permite identificar condições que não são expressas apenas por uma ou duas variáveis, mas se constituem condições complexas (variáveis latentes), resultante da atuação de um conjunto de variáveis e permite identificar também variáveis mediadoras. A MEE pode contribuir para explicitar as relações existentes entre as variáveis que estão, possivelmente, no caminho causal.

Será realizada uma análise para se identificar a influência de variáveis complexas como as condições socioeconômicas, aceitação da gestação e verificar se a assistência pré-natal exerce papel de mediação entre estas variáveis e o óbito fetal anteparto. Para a realização deste estudo, serão utilizados os dados da pesquisa “Mortalidade perinatal na região sul do município de São Paulo: um caso controle de base populacional” (Almeida et al. 2006).

Foi desenhado um diagrama do possível caminho causal para óbitos fetais, a partir do quadro conceitual proposto inicialmente em Almeida et al. (2007), em que se pressupõem relações causais entre as variáveis de estudo, partindo-se de variáveis distais (socioeconômicas e da mãe) e passando pelas proximais (biológicas). O diagrama apresenta o modelo que expressa a relação entre as variáveis latentes SSE (que representa a situação socioeconômica) e nAdG (que representa a não-aceitação da gestação), e as demais variáveis em estudo e o desfecho (Figura 1).

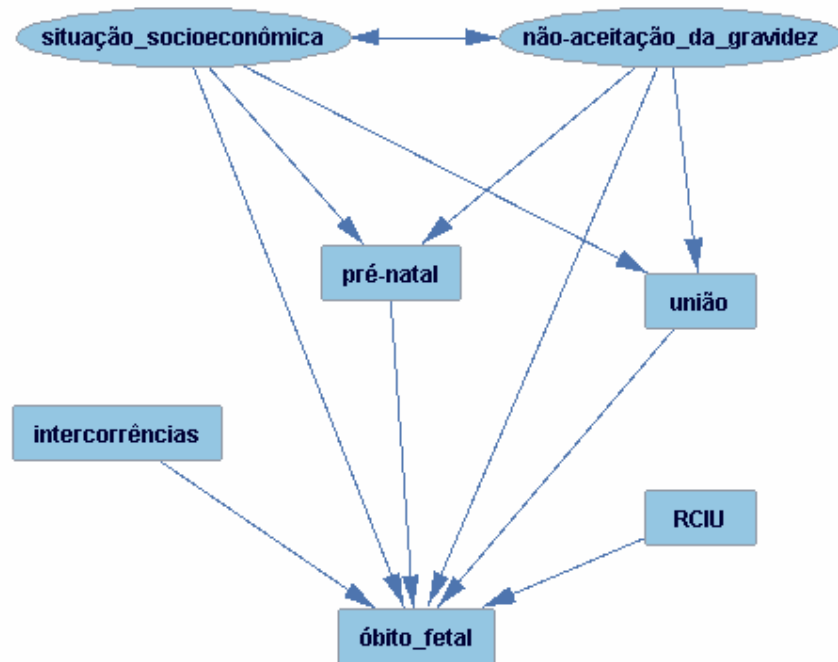


Figura 1. Caminho causal para o óbito fetal.

1.2 Multicausalidade

O conceito de risco, em epidemiologia, é definido como a probabilidade de uma doença se desenvolver em um indivíduo em um intervalo de tempo (MacMahon 1970, Susser 1973, Rothman 1998).

Rothman (2002) retoma a concepção de risco indicando que a doença pode ter mais de um mecanismo causal ou que “diferentes mecanismos de exposição” são formados por combinações de fatores provindos de diferentes dimensões como, por exemplo, variáveis de meio ambiente, de inserção socioeconômica, de complicações médicas etc.

Diversos autores indicam a importância de um quadro teórico-conceitual (Victora 1997, Rothman 1998) para estabelecer as possíveis relações existentes entre as diferentes exposições em estudo e o desfecho. A construção de um quadro conceitual para cada grupo de variáveis e seus componentes depende do conhecimento científico prévio das possíveis relações de diferentes exposições (fatores de risco), que possam estar associados ao desfecho estudado e a sua concepção deve ser feita no início do estudo, antes de definir seu delineamento. Na construção do quadro, a temporalidade é um item importante, pois se considera que a causa de uma doença é um evento, condição ou característica que precede a doença e, sem ela, a doença não teria ocorrido de forma completa ou poderia ocorrer posteriormente no tempo (Hill 1971).

O quadro referencial teórico é uma tentativa de aproximação da realidade, pois busca estabelecer o conhecimento a respeito das diversas exposições possivelmente presentes em determinada doença. Portanto, é necessário identificar as possíveis relações entre estas exposições e o desfecho do ponto de vista epidemiológico. Paralelamente, as informações a respeito das variáveis que representam essas possíveis relações são utilizadas para se definir o delineamento do estudo. A partir desse delineamento, define-se o modelo estatístico para a análise e, ao se obter os dados, é feito o tratamento numérico de fatores relevantes (exposições/construtos), de variáveis aleatórias, estimação de parâmetros populacionais e testes estatísticos de uma ou mais hipóteses (Kleinbaum et al. 1982).

A modelagem fornece estimativas dos parâmetros para as associações entre as diferentes exposições e o desfecho. Por exemplo, em estudos de caso-controle, a medida que quantifica a magnitude dessa associação é a razão de odds (*odds-ratio*). A razão de odds é definida como sendo a razão entre a chance de o desfecho ocorrer em um grupo exposto e a chance de o desfecho ocorrer em outro grupo exposto. Chance é a probabilidade de ocorrência de um evento dividida pela probabilidade de não ocorrência deste mesmo evento (Gordis 2004).

A construção metodológica de um estudo é orientada por diversas hipóteses levantadas previamente. Cada hipótese é uma tentativa de explicação do problema como, por exemplo, quando se faz a pergunta se determinado fator é importante para explicar o desfecho estudado. Essa questão pode ser escrita de forma matemática que irá confrontar a hipótese considerada pelo pesquisador com sua negação (hipótese nula). A hipótese considerada pressupõe a existência de diferença (ou somente maior ou somente menor) entre os grupos ou subgrupos estudados em relação a algum parâmetro, como a média de uma variável, enquanto a hipótese nula nega que exista tal diferença. O objetivo do teste estatístico de hipóteses é permitir a verificação se os dados da amostra trazem evidência para apoiar ou não uma hipótese estatística formulada (Bussab e Morettin 2002).

A modelagem estatística é baseada em testes de hipóteses que consideram níveis de significância (α), fixados à priori ou, alternativamente, fornecendo níveis descritivos (p-valores). A estatística de teste é calculada a partir dos dados e o teste estatístico é conduzido admitindo-se uma distribuição probabilística, sob a hipótese nula. Os resultados podem ser expressos a partir do valor de α para se construir uma regra de decisão ou, de outra forma, ser expressos em termos do nível descritivo (p-valor) que representa a probabilidade de se obter estimativas mais desfavoráveis do que está sendo fornecida pela amostra, supondo a hipótese nula verdadeira (Magalhães e Lima 2002).

Considerando que as várias exposições podem estar envolvidas em determinado desfecho, avaliam-se esses efeitos conjuntamente. A partir daí, gera-se um modelo estatístico de teste para diversas hipóteses a respeito do problema estudado (Harrell 2001). Entretanto, a inferência estatística não é o mesmo que inferência causal, pois

ambas as inferências tratam de objetos diferentes (Luiz e Struchiner 2002). Para esses autores, a Epidemiologia busca identificar a etiologia das doenças e utiliza a estatística no processo de inferir causalidade; a Estatística, por sua vez, interpreta o problema de atribuição de causalidade como um problema de detectar efeitos causais por meio de mensurações. Assim, geralmente, a inferência estatística é utilizada na avaliação de dados para seu uso na inferência causal. Uma discussão presente em textos de epidemiologia (Rothman 1998, Szklo 2000) é que a exclusão de uma variável não deve ser feita apenas por critério estatístico, uma vez que a explicação baseada somente em uma medida de probabilidade não estabelece a causalidade, dado que existem outras possíveis razões não causais para a existência dessa associação.

Para analisar as diversas causas e diversos caminhos possíveis da ocorrência de um evento, normalmente utilizam-se modelos estatísticos que permitem compreender a relação entre variáveis explicativas e o desfecho, tornando possíveis os testes dos modelos delineados pelo pesquisador. O quadro conceitual expressa um dos caminhos possíveis das relações entre as variáveis e o desfecho, e este pode ser utilizado para guiar a análise, na medida em que define a ordem de entrada de variáveis (ou pelo menos dos conjuntos, ou blocos de variáveis) no modelo. Diversas técnicas analíticas mais simples e conhecidas são usualmente aplicadas para obter estimativas das relações entre as variáveis de diversas dimensões de exposição e o desfecho estudado. A análise de modelos de regressão múltipla é a mais utilizada na literatura em epidemiologia na qual é medida a relação entre cada variável e o desfecho e estimativas de seus efeitos são obtidas conjuntamente na presença de outras variáveis no modelo. Os desfechos binários são bastante comuns, como morte/não-morte; doença/não-doença e, para esse tipo de variável, são utilizadas algumas funções de modelagem além da logística, como as funções proibito, complemento log-log e log-log (McCullagh 1989).

A função logística é utilizada principalmente por fornecer diretamente uma estimativa do risco, como a razão de odds. A razão de odds é uma estimativa da associação entre a exposição e o desfecho em estudos de caso-controle. A partir dos coeficientes dessa regressão, estimativas de razão de odds e de seus erros-padrão são

obtidos os testes que auxiliam na identificação de variáveis importantes no modelo. Contudo, os testes são realizados sobre essas estimativas e não refletem necessariamente causalidade, como já foi extensivamente discutido (Szklo 2000, Gordis 2004).

Modelos estatísticos desenvolvidos para entender as relações causais, são chamados de modelos causais. Estes modelos utilizam a lógica de causa e efeito para descrever a realidade. Além disso, buscam a associação de variáveis em estudos a fatores de natureza econômica ou política. Segundo Greenland e Brumback (2002), na área da saúde, quatro são os modelos causais: diagramas causais ou modelos por grafos (*causal diagrams* ou *graphical causal models*), modelos “contrafactuais” ou de “desfecho potencial” (*counterfactual models* ou *potencial-outcome models*), modelos de causas suficiente-componente (*sufficient-component cause models*) e modelos de equações estruturais (*structural equations models*). Os modelos por grafos e os de desfecho potencial são usados para descrever a presença de interações causais, porém não seu mecanismo. Os outros dois pretendem estudar o mecanismo causal. Vale mencionar que Pearl (2000) mostra que o modelo chamado “modelo causal de Rubin” (*Rubin causal model* –, Rubin 1974) é equivalente ao modelo de equações estruturais.

1.3 Variável de confusão e variável mediadora (interveniente)

Dentre as diversas exposições que podem estar presentes em determinado desfecho, algumas são consideradas como variáveis explicativas e, destas, um conjunto importante é formado pelas variáveis de confusão, que podem alterar as relações de interesse e devem ser incluídas no modelo de análise. Segundo Szklo (2000), para se considerar uma variável de exposição como sendo de confusão, ela deve estar associada de forma causal com o desfecho e com outra variável de exposição (de forma causal ou não-causal).

As variáveis de confusão têm que ser cuidadosamente trabalhadas para evitar problemas nas estimativas de efeitos e seus erros-padrão (Hosmer e Lemeshow

1989). Existem algoritmos para a identificação de variáveis de confusão já descritos (Rothman 1998; Greenland 1989).

Um tipo de variável importante é a variável mediadora. A variável mediadora (ou interveniente) é uma variável de exposição que pode ser explicada por uma outra variável de exposição e que, por sua vez, tem um efeito sobre o desfecho. Assim, a variável mediadora se coloca entre a variável explicativa e o desfecho, de forma que ela altera o desfecho e é alterada pela variável explicativa. Quando a exposição determina parte da presença ou da expressão da possível variável de confusão, que também está associada ao desfecho, tem-se uma variável “intermediária”, que faz a ligação entre a exposição e o desfecho (Szklo 2000).

A distinção de variáveis de confusão das variáveis chamadas intervenientes ou mediadoras deve ser explicitada nos estudos epidemiológicos (Spencer 2005, Olinto 1998, Susser 1973).

Susser (1973) identifica a presença de uma variável interveniente no seguinte exemplo: o peso ao nascer de uma criança está associado à dieta materna desde o início da gravidez e um pesquisador poderia considerá-la como sua causa. Por outro lado, pode-se imaginar que a dieta da mãe é uma variável antecedente no tempo e que ela teria determinado o peso da mãe. Segundo o autor, estudos indicam que parte do ganho de peso materno durante a gestação se dá principalmente até 20 semanas de gestação e que esse ganho de peso está associado com o peso ao nascer do recém-nascido. Assim, a dieta da mãe pode ser uma variável causal que antecede e determina o peso da mãe até a 20^a semana de gestação (isto é, não é o peso que determina a dieta). Enfim, a variável peso da mãe é uma variável interveniente (Figura 2). Susser (1973) afirma que quando uma variável interveniente é controlada na análise, a associação inicial entre a variável explicativa e o desfecho pode desaparecer ou ter seu efeito (risco) reduzido.

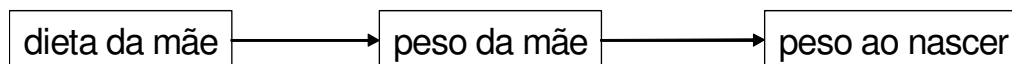


Figura 2. Peso da mãe como variável mediadora entre dieta e peso ao nascer da criança.

Em um estudo metodológico dessas variáveis, MacKinnon et al. (2002) utilizam como sinônimos os termos “mediação” e “interveniência”. Diversos trabalhos também têm utilizado os termos “mediadora” e “interveniente” como sinônimos, ao se tratar essas variáveis. Por exemplo, Brook et al. (1995) empregam o termo processo “interveniente” no título do trabalho e indicam que o caminho entre a agressão infantil e o uso de drogas nas idades de 15 a 20 anos, foi “mediado” pelo estresse psíquico.

Em pesquisa na base de dados do *International Journal of Epidemiology*, o termo mais frequentemente utilizado é “mediação”. Esta denominação foi encontrada nos seguintes trabalhos: Blakely et al. (2003) apresentaram a renda como variável mediadora entre a educação materna e a mortalidade infantil; Breslau et al. (2005) apresentaram o baixo peso ao nascer como variável mediadora para a associação existente entre o hábito de fumar da mãe e baixo QI de crianças com idades de 6, 11 e 17 anos; Kilander et al. (2001) identificaram mudanças no estilo de vida como variável mediadora que inverte a relação existente entre baixo nível de escolaridade e morte por doença cerebrovascular; Osler et al. (2003) estudaram variáveis mediadoras na relação existente entre renda e doença isquêmica do coração. Forssas et al. (1999) indica o baixo peso ao nascer como variável mediadora da relação existente entre condições socioeconômicas e a mortalidade perinatal.

Szklo (2000) alerta que uma variável de confusão não deve ser considerada como variável mediadora quando esta não se encontra no caminho causal entre a exposição e o desfecho. Assim, não é apropriado realizar o ajuste das medidas de efeito das demais variáveis em estudo por uma variável mediadora, a não ser que o pesquisador tenha como objetivo identificar mecanismos de explicação entre as exposições e o desfecho (Szklo 2000). Assim, a variável de mediação não deve ser incluída no modelo em que se quer avaliar o efeito das diversas exposições sobre o desfecho, pois a pergunta de interesse é “determinada exposição causa o desfecho?”. Entretanto, uma vez que se sabe ou se estabelece a relação entre a variável mediadora e o desfecho, é possível avaliar se a exposição causa o desfecho por algum outro mecanismo que não passe somente pela variável mediadora (Szklo 2000) e introduzir a variável mediadora no modelo.

1.4 Modelos de Equações Estruturais (MEE)

Os modelos de equações estruturais (MEE) incluem vários procedimentos tradicionais de análise de dados, como a regressão e a análise fatorial, além de incluir outros menos conhecidos, como a análise de caminhos (*path analysis*) que objetiva analisar relações causais entre variáveis observadas (Vasconcelos et al. 1998) e o MEE com variáveis latentes.

Dois conceitos importantes são utilizados nos MEE: os de variável **observada** - (ou manifestável) é aquela que sua ocorrência é observada, como o peso de um bebê, por exemplo - e de variável **latente** (não-observada). Para Skrondal (2004), “variável latente é uma variável aleatória cuja ocorrência está escondida”. O significado da variável latente pode variar: 1) a variável estudada é medida com erro, ou seja, tem-se uma variável observada, mas há um erro associado à medida tomada, a variável latente representará a verdadeira medida da variável (aqui, presume-se que ela existe); 2) a variável latente pode ainda representar construtos hipotéticos que contemplam um aspecto epistemológico ou, segundo Crombach, um “aparato intelectual” que representa conceitos; 3) a variável latente pode ser vista, também, como representando o efeito de um conjunto de variáveis que podem ser medidas, mas que não o foram como, por exemplo, em um modelo de regressão em que o termo de erro pode representar variáveis não medidas. Entretanto, a essas variáveis não se atribuem um significado como se faz com um construto; 4) variáveis latentes podem representar variáveis existentes, porém com valores faltantes que, por algum motivo, não se conseguiu registrar a ocorrência; 5) variáveis latentes podem representar variáveis contínuas que, no entanto, foram medidas de forma dicotômica ou ordinal.

Modelos de equações estruturais são formados por duas partes: o **modelo de mensuração** (*measurement model*), na qual as variáveis latentes são construídas a partir de variáveis observadas, e o **modelo estrutural** (*structural model*), que trata da relação entre as variáveis latentes. Outras variáveis de exposição podem ser incluídas

em modelos de regressão utilizando as variáveis latentes como variáveis explicativas e/ou como variáveis respostas.

Há outras denominações para as variáveis em textos de econometria sobre modelos de equações simultâneas em que aparece a denominação de variável endógena e exógena e é encontrada em alguns livros sobre MEE. As variáveis **endógenas** são aquelas que são determinadas pelo sistema de equações, em função de outras variáveis em estudo. As variáveis **exógenas** são determinadas por fatores externos ao sistema e, em regressão múltipla, podem ser entendidas como variáveis explicativas. Esses termos são utilizados em alguns livros sobre MEE (Klem 1995, Kaplan 2000).

No MEE, o **diagrama de caminhos** (*path diagram*) irá explicitar as hipóteses teóricas do pesquisador sobre as relações existentes entre os diversos níveis de exposição e das variáveis. Historicamente, o primeiro modelo a estabelecer o uso de caminhos foi o modelo de equações simultâneas (*path analysis*) proposto por Sewall Wright, em 1921, que, em sua análise de trajetórias, utilizava modelos de regressão para estudar o efeito de variáveis. Wright denominava essas variáveis de variáveis antecedentes no estudo de morfologia animal (Kaplan 2000). A econometria também utiliza equações simultâneas em modelagem de ocorrências no campo macroeconômico para dados longitudinais, em estudos de painel (Wooldridge 2003).

1.4.1 Representação gráfica - diagrama de caminhos

O **diagrama de caminhos** é uma representação gráfica das relações entre as variáveis observadas e latentes em que a direcionalidade é um item importante e é representada por meio de setas que indicam, teoricamente, a direção de causa e de correlação entre as variáveis. Para apresentar os elementos dos diagramas, será utilizado o termo influência para indicar as relações entre as variáveis (Hair et al. 2005). Nos diagramas de caminhos (Figura 3), as variáveis observadas são representadas por retângulos e cada *caminho* é representado por uma linha com uma seta em uma das extremidades. As variáveis explicativas podem ser ligadas por curvas com setas nas duas extremidades que representam *correlações* existentes entre os erros destas variáveis. A seta apontada para a variável resposta é o do termo

de erro da variável, ou erro do resíduo. Eventualmente, o termo de erro, representado aqui por ε , vem dentro de uma oval. Essa representação também se aplica aos modelos de equações estruturais com variáveis latentes que será apresentado no item a seguir.

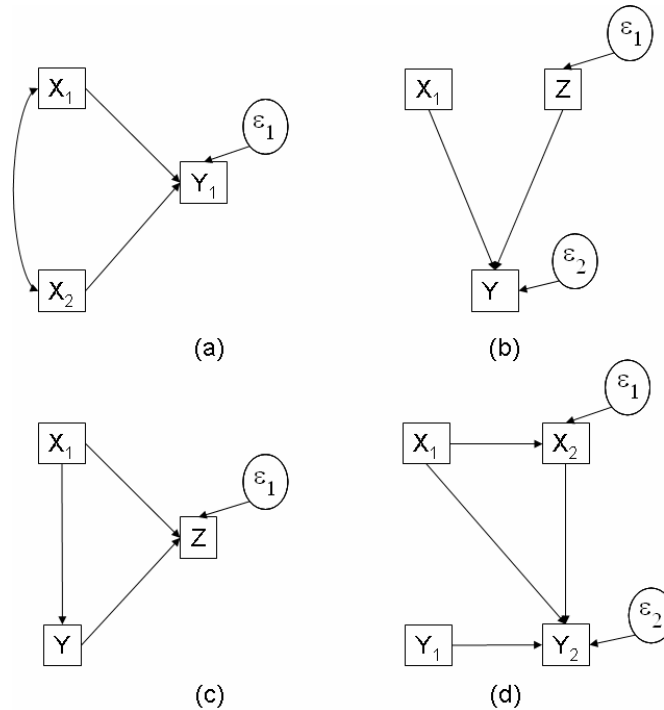


Figura 3. Quatro diagramas de caminhos.

A Figura 3(a) representa um modelo de regressão múltipla com duas variáveis exógenas (X_1 e X_2) e uma endógena (Y) e a seta curva indica correlação entre X_1 e X_2 . Figura 3(b) representa um modelo de *análise de caminhos* com *mediação*, em que Y modifica o efeito de X em Z . Nesse caso, pode-se imaginar que não há efeito *direto* de X_1 em Z , há uma mediação, ou interveniência, promovida pela variável Y . Bollen (1989) denomina esta situação de mediação completa. A Figura 3(c) é mais interessante, pois combina *a* e *b*, em que X_1 tem um efeito direto sobre Z , mas também atua sobre Y , que por sua vez afeta Z . Na situação da Figura 3(d), X_1 é um fator que influencia X_2 e de forma semelhante Y_1 influencia Y_2 . E ainda X_1 e X_2 influenciam Y_2 . Nesse caso, pode-se admitir que X_1 e X_2 são variáveis antecedentes temporalmente assim como Y_1 é antecedente a Y_2 . A temporalidade pode ser inserida nessa representação e tem um papel muito importante. O exemplo da Figura 3(d) representa situações do efeito de diferentes variáveis ocorridas no mesmo tempo, por

exemplo, X_1 representa uma medida de ansiedade da mãe e Y_1 representa uma medida de depressão da mãe e X_2, Y_2 são medidas do comportamento de seu filho. As variáveis X_1 e Y_1 são teoricamente antecedentes às medidas utilizadas que representam o comportamento dos filhos. Desse modo, está-se levando em conta um efeito temporal.

1.4.2 A análise fatorial no contexto de MEE

O MEE possibilita a estimação do modelo de mensuração (que pode ser uma análise fatorial confirmatória ou um modelo de regressão com resposta latente) e especificar o modelo estrutural (Hair et al. 2005). Assim, o MEE é flexível para se trabalhar com equações relacionadas entre si simultaneamente e fornece uma visão detalhada das possíveis relações causais entre os construtos. As variáveis que definem a variável latente (construto) no modelo de mensuração são denominadas de indicadores.

A descrição matemática do MEE com seus componentes se encontra no Anexo 1, além de apresentar alguns casos particulares de MEE, tais como a regressão linear, análise de caminhos (*path analysis*), análise fatorial e MEE com variáveis latentes.

Apresentam-se a seguir informações sobre a análise fatorial e modelos de equações estruturais com variáveis latentes.

Análise fatorial

A análise fatorial parte da idéia do emprego de variáveis latentes. Na análise fatorial, as variáveis latentes representam construtos não-observados, que são chamados de fatores ou dimensões. Como resultado da análise fatorial, tem-se cargas fatoriais e escores fatoriais. Uma carga fatorial representa a correlação entre uma variável e o fator considerado, de forma que quanto maior for a carga de uma variável observada, mais a variável se identifica com o que se está considerando como fator. Os escores fatoriais são os valores estimados dos fatores comuns que são calculadas para cada indivíduo e pode ser utilizado como uma entrada para análises posteriores. Neste tipo de análise, muitos estudos de tipo ecológico têm empregado a técnica de componentes principais (Dunteman 1989) para agrupar fatores semelhantes

(Koifman 1997). Koifman et al. (1997) utilizaram o método de componentes principais para identificar a associação entre câncer de estômago e o consumo de alimentos enlatados.

Uma possibilidade de se trabalhar com o resultado da análise fatorial é utilizar os escores fatoriais dos construtos para serem utilizados em modelos de regressão com as outras variáveis do estudo. Os escores fatoriais gerados são variáveis contínuas e podem ser utilizados diversos tipos de regressão, como a linear ou a logística, por exemplo. O uso dessa abordagem não é uma solução simples de ser aplicada em casos de modelos mais complexos, como no MEE, em que pode existir correlação relevante entre os indicadores (usados para construir o fator) e outras variáveis observadas que sejam de interesse no quadro conceitual (Muthén 1984). No MEE, ao se utilizar matrizes de correlação, as relações podem ser feitas dentro de um mesmo modelo e o resultado é um estudo de como os construtos se comportam entre si e em relação às outras variáveis do estudo.

No MEE, é necessário informar os parâmetros que serão estimados. Cada modelo é único frente à hipótese levantada, o que diferencia da análise fatorial exploratória em que o modelo é desenhado para se tentar determinar a composição das variáveis latentes.

1.4.3 Limitações da análise

A falta de identificabilidade em MEE é um problema na estimação que leva a diversos erros na prática. Esse é um assunto complexo e não existe uma regra geral para identificar o problema. Detalhes sobre problemas de identificação estão no Anexo 2. Klem (1995) indica que, ao se analisar os resultados de uma análise, o pesquisador deve considerar a possibilidade da existência dos seguintes problemas: erros de medidas nas variáveis observadas; presença de erros de especificações no modelo; e presença de multicolinearidade.

A presença de determinada multicolinearidade nos modelos de mensuração, pelo fato de existir correlação entre as variáveis, é bastante comum e pode não ser visto como

um problema. Na análise fatorial, a existência de colinearidade ajuda a agrupar as variáveis em indicadores para a construção das variáveis latentes (Bollen e Lennox 1991). Em modelos de regressão, a multicolinearidade pode ser um problema em casos de alta correlação, pois isso pode provocar uma inflação nos erros-padrão das estimativas (Neter et al. 1996).

Algumas limitações são mencionadas por Bollen (1989) ao se utilizar os MEE: a) a necessidade do uso de variáveis latentes; b) a presença de suposições sobre a distribuição das variáveis latentes; e c) falseabilidade do modelo ao considerá-lo como sendo causal. Com referência a essa última limitação, Yu (2007) cita Ling (1982) que afirma que o pesquisador nunca vai negar uma falsa suposição causal, enquanto que Freedman (1997) argumenta que o caminho causal traçado não é derivado apenas da teoria causal formulada a partir de dados empíricos, pois o modelo inicial é definido pelo pesquisador, com base no conhecimento sobre o tema, fato que é precedente à análise dos dados.

Quanto ao tamanho de amostra necessário para o emprego de MEE, existem várias regras (tais como 5 a 10 observações por parâmetro, 50 observações por variável e não menos do que 100 indivíduos) (Muthén e Muthén 2002), ou, conforme Hox e Bechger (1998), “pelo menos 500 a 1000 indivíduos”. O tamanho da amostra necessária depende do número de variáveis, ocorrência de valores faltantes (*missing*), confiabilidade das variáveis e a força da relação existente entre as variáveis. Em um modelo de simulação de análise fatorial confirmatória, em que se empregam 2 variáveis latentes contínuas e 5 indicadores contínuos cada, foram obtidos 150 indivíduos para um poder de 0,81 para rejeitar a hipótese de que a correlação do fator é zero. Para indicadores não-normais, 265 indivíduos para o poder de 0,80 (Muthén e Muthén 2002).

1.4.4 O MEE e as variáveis categóricas ordinais e binárias – estimação

No contexto de modelos de equações estruturais, há alguns pontos que necessitam atenção quando se empregam variáveis categóricas, como é o caso do presente estudo.

Usualmente, os MEE são baseados na suposição de que as variáveis medidas (indicadores) apresentam uma dada estrutura de covariância, e têm uma distribuição normal e são identicamente distribuídas (Bollen 1989). Assim, a estimação dos parâmetros de MEE pode ser feita via métodos de máxima verossimilhança (Jöreskog 1973), ou o de mínimos quadrados generalizados (*generalized least squares*, GLS) - utilizado em MEE por Jöreskog e Goldenberg (1972) e Browne (1974) -, e o método de mínimos quadrados ou ordinários (ULS).

Variáveis observadas não-normais podem ser variáveis contínuas (com assimetria e curtoses mais acentuadas com relação à distribuição normal) ou variáveis categóricas especialmente com poucas categorias ou binárias. Bollen (1989) sugere que este pressuposto sempre é violado, devido aos limites dos instrumentos de medida. Este não é um problema sério quando se tem uma variável categorizada em diversas categorias como, por exemplo, em uma escala de 0 a 10.

A utilização de variáveis categóricas ordinais ou de variáveis que apresentam poucas categorias ou ainda variáveis binárias, comuns em estudos epidemiológicos, tem conseqüências sobre o pressuposto do modelo de mensuração de Y , $y = \Lambda_y \eta + \varepsilon$, em que y deve ter distribuição normal. Isso também serve para X , na notação apresentada. Quando as variáveis observadas (medidas) no estudo são categóricas, essa pressuposição não ocorre. Essa situação também interfere nas covariâncias entre X e Y , ao se pressupor que X e Y tenham distribuição normal bivariada (Bollen 1989).

Alguns métodos para se tratar com variáveis não-normais são: a) estimadores de distribuição assintótica livre (ADF); b) ajuste do qui-quadrado da estimação por máxima verossimilhança; c) utilizar uma versão robusta para o método de mínimos quadrados ponderados (*weighted least squares*, WLS) (Muthén 2004); d) utilizar algoritmo *bootstrap* para cada parâmetro e a estatística qui-quadrado para obter erros-padrão mais acurados e p-valores associados ao qui-quadrado (Finney e Disteffano 2006).

No MEE, para evitar o problema da não-normalidade das variáveis categóricas, é necessário o uso de um modelo de pontos-de-corte ou **limiares** (*threshold model*). Para se estimar corretamente a matriz de covariâncias de indicadores dicotômicos ou

ordinais, assume-se que essas variáveis observadas derivem de variáveis contínuas que tenham distribuição normal. Assim, estima-se a correlação entre as variáveis da seguinte forma: se as variáveis são categóricas, a correlação entre as variáveis latentes correspondentes é chamada de policórica; a correlação entre duas variáveis dicotômicas é tetracórica e se uma variável for ordinal e a outra contínua, ela é poliserial (Bollen 1989). Jöreskog & Sörbom (2002) recomendam que, se os dados são ordinais ou categóricos, o método de estimação a ser empregado é o de mínimos quadrados ponderados (WLS) e a matriz de correlação utilizada deve ser a tetracórica ou policórica. Esses métodos estão detalhados no Anexo 3.

Assim, a variável categórica y é considerada como proveniente de uma outra variável, chamada y^* (não medida) que seja contínua e tem distribuição normal com média 0 e variância 1. Os limiares serão os pontos-de-corte de y^* e representam as categorias de y . Por exemplo, se uma variável observada tem 3 categorias (1, 2 e 3), pode-se supor que existe uma variável, como a da Figura 4, em que os limiares k_1 e k_2 representam pontos-de-corte de forma que: se a variável $y^* < k_1$, então tem-se que y é igual a 1, se y^* estiver entre k_1 e k_2 , y é igual a 2; e se $y^* > k_2$, então y é igual a 3.

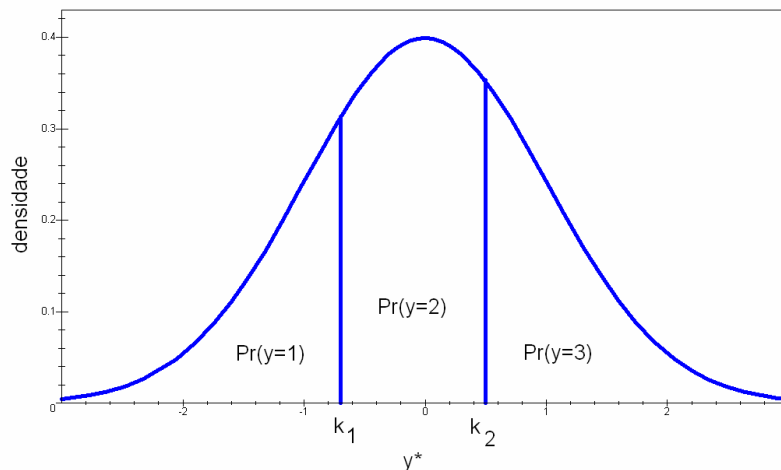


Figura 4. exemplo de variável y^*

Em resumo, calculam-se as correlações policóricas, poliserial ou tetracórica para os pares de variáveis com distribuição conjunta não-normal, assumindo-se que eles são gerados por uma variável latente contínua, que tenha sua distribuição conjunta (*large sample joint distribution*) bivariada normal. Essas correlações são inseridas na estimação.

O método de estimação que será utilizado na estimação na análise da mortalidade fetal é o proposto por Muthén (1984) que apresenta os seguintes passos: a) estimar os pontos de corte ou limiares (*thresholds*) populacionais; b) estimar as correlações (policórica, tetracórica) populacionais considerando os pontos de corte; c) estimação dos parâmetros do modelo, com base nas correlações e pontos de corte obtidos. A descrição mais detalhada dessa estimação está no Anexo 3.

1.4.5 Etapas do MEE

James et al. (1982) propuseram a modelagem em duas etapas: primeiro, estabelecer o modelo de mensuração para todas as variáveis latentes e, então, proceder ao modelo estrutural. Mulaik e Millsap (2000) sugerem quatro etapas: (a) um modelo fatorial, (b) um modelo fatorial confirmatório, (c) um modelo de equações estruturais, d) possivelmente, um modelo com mais restrições nos parâmetros do modelo para simplificá-lo. Essa abordagem é criticada por Hayduk e Glaser (2000).

A análise fatorial exploratória é aplicada na situação em que não se conhece com mais detalhes a teoria a que se refere o evento estudado. Por outro lado, a análise fatorial confirmatória pode ser utilizada quando se conhece a teoria a respeito das relações entre as variáveis e os fatores para determinado estudo.

A combinação de análise fatorial exploratória e depois a análise fatorial confirmatória é uma opção quando se tem uma teoria subjacente, mas não se conhece, a partir das variáveis medidas, o número de fatores ou mesmo se há necessidade de verificar sua composição (Hatcher 1994). Este autor sugere, no entanto, fazer a análise fatorial exploratória em metade dos dados e a análise fatorial confirmatória com a outra metade dos dados, desde que o tamanho da amostra seja razoável para isso. Se houver problemas no ajuste da análise confirmatória, pode-se fazer uma análise exploratória para identificar inconsistências entre os dados e o modelo, mas a aplicação da análise confirmatória a partir das alterações propostas deveria ser feita em um novo conjunto de dados. A utilização da análise fatorial exploratória em dois subconjuntos como uma primeira etapa da análise para verificar se o mesmo número de fatores se ajusta bem a cada subconjunto parece adequada. A

utilização da análise confirmatória em todo o banco de dados serve como alternativa para identificar quais são as variáveis observadas associadas a determinado(s) fator(es) - trabalha-se a análise fatorial exploratória no contexto de uma análise confirmatória, avaliando-se a necessidade de se identificar variáveis observadas que podem compor mais de uma variável latente, além de identificar a necessidade de incluir relações dos erros entre variáveis observadas de um mesmo fator ou de fatores diferentes (Jöreskog 1979).

1.4.6 Interpretação

O resultado principal de um MEE está na observação das relações de influências entre as variáveis do modelo, cuja significância de cada relação e ajuste do modelo são obtidos. Em análises com o MEE, usualmente são analisados as medidas a partir dos coeficientes calculados a partir das variáveis padronizadas para se retirar o efeito da magnitude das variáveis e permitir a comparação entre as variáveis do modelo.

Para a parte do modelo de mensuração (análise fatorial), o coeficiente não-padronizado representa o número de unidades que se altera na variável observada (indicador do fator) por unidade que se altera na variável latente (score fatorial). Como o score fatorial não tem dimensão, o valor (número) desses coeficientes não tem interpretação prática, porém o sinal do coeficiente passa a ser importante, pois a variação (aumento ou diminuição) do score representa a variação (dada pelo sinal) da variável que compõe o fator. Se um coeficiente do modelo estrutural (entre as variáveis latentes) é igual a 2,0, então a variável latente resposta apresenta um aumento de 2,0 unidades padronizadas para cada aumento de uma unidade da variável latente explicativa.

Para a relação entre as variáveis latentes e as variáveis binárias que são resposta do modelo, os coeficientes são os de uma regressão probito. Usualmente, somente seus sinais e significâncias são utilizados. Um valor positivo desse coeficiente significa que a probabilidade de uma variável categórica dependente (por exemplo, para a categoria 1 de uma variável 0/1) é aumentada quando a variável explicativa aumenta. Maiores valores significa que esta probabilidade aumenta mais rapidamente.

Uma vantagem dos MEE é a possibilidade de se separar o efeito total das variáveis explicativas em dois efeitos: o efeito direto, que expressa a relação direta de uma variável sobre outra, e o efeito indireto, na qual o efeito de uma variável se exerce intermediada por uma terceira variável (mediadora) (Bollen 1989). Isso pode auxiliar na melhor compreensão das relações entre as variáveis. A figura 5 ilustra a regressão entre a exposição A, a mediadora B e desfecho C: supõe-se que a variável A tem efeito direto sobre C (não representado na figura) dado por $C = \beta A$. Decompondo-se β em uma soma do efeito direto de A em C e do efeito indireto de A em C via B, podemos escrever: $B = \gamma_2 A$ e $C = \gamma_1 A + \gamma_3 B$. O efeito total de A é calculado como a soma do efeito direto de A, somado à multiplicação dos efeitos indiretos de A ($\gamma_1 + \gamma_2 \times \gamma_3$).

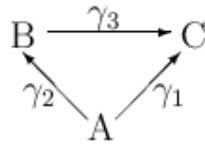


Figura 5. exemplo de mediação.

2 OBJETIVOS

Objetivo geral

Este estudo objetiva estudar o efeito da situação socioeconômica, da não-aceitação da gravidez e da assistência pré-natal na mortalidade fetal anteparto, na região sul do município de São Paulo, em 2000.

Objetivos específicos

1. Obter estimativas e avaliar o efeito dos construtos “situação socioeconômica” e “não-aceitação da gravidez” na mortalidade fetal.
2. Avaliar o papel da assistência pré-natal na mortalidade fetal.
3. Mostrar, passo-a-passo, a respectiva modelagem de equações estruturais para análise da mortalidade fetal, na presença de variáveis observadas binárias e contínuas.

3 METODOLOGIA

3.1 Material

3.1.1 Fonte de dados

Os dados foram obtidos da pesquisa “Mortalidade perinatal na região sul do município de São Paulo: um caso controle de base populacional”, realizada em 2000. Estudos iniciais sugeriram que os fatores de riscos pudessem ser diferentes para os dois componentes da mortalidade perinatal – óbito fetal e neonatal precoce, o que levou a se considerar desfechos separados (Almeida et al. 2007). No presente trabalho, foi estudada apenas o componente de mortalidade fetal.

3.1.2 Tipo de estudo

O tipo de estudo da pesquisa é caso-controle e contou com 164 casos representados pelos óbitos fetais anteparto e 313 controles representados por recém-nascidos ao período neonatal (27 dias ou mais). A pesquisa foi delineada para se estudar a mortalidade perinatal (óbitos fetais e neonatais precoces) em que foram analisados 164 óbitos fetais e 146 óbitos neonatais que resultou em um total de 310 casos de óbitos perinatais e 313 controles (Almeida et al. 2007).

3.1.3 Período de referência e população de estudo

A população de estudo corresponde aos 313 sobreviventes ao período neonatal e 164 óbitos fetais anteparto de mulheres residindo na região sul do município de São Paulo, no período entre 1º de agosto de 2000 a 31 de janeiro de 2001. O estudo foi conduzido em 14 distritos localizados na região sul do município de São Paulo. Estes distritos contam com aproximadamente 23% da população e ocupam 44% da área da cidade (Almeida et al. 2007).

3.1.4 Variáveis de estudo

As variáveis observadas contínuas utilizadas nas análises são:

- anos de escolaridade do chefe de família;
- anos de escolaridade da mãe;
- renda familiar per capita (em salários mínimos);

As variáveis observadas categóricas utilizadas nas análises são:

- união (sem união ou com união no momento da gravidez);
- tipo de habitação (alvenaria, demais);
- classificação socioeconômica (A, B, C, D ou E), baseado em pontos atribuídos à escolaridade da mãe e do pai e da posse de bens familiar (Critério de Classificação Econômica Brasil - critério Brasil), produzida pela Associação Brasileira de Empresas de Pesquisa, e utilizada pela Fundação SEADE (ABEP 2000);
- densidade pessoa-cômodo (número de pessoas por cômodo, excetuando-se banheiro);
- ocupação do chefe de família (manual, qualificado, sem emprego), baseado na classificação britânica de ocupação (Borrell 2003);
- plano de saúde (sim, não);
- planejamento da gestação (sim, não);
- a mãe pensou em aborto durante a gestação atual? fez algo para abortar? (não, sim e não tentou, sim e tentou);
- reação negativa da mãe, do pai e da família, quando soube da gravidez (feliz, infeliz/indiferente);

- intercorrências durante a gestação (sim, não): se a mãe apresentou alguma das seguintes intercorrências: sangramento (sim, não), estado hipertensivo (sim, não), diabetes (sim, não), doença renal (sim, não);
- adequação do cuidado pré-natal (adequado, inadequado que inclui a ausência de pré-natal). O pré-natal foi considerado adequado quando a primeira consulta ocorreu no primeiro trimestre da gestação, com no mínimo quatro consultas e em pelo menos uma das consultas foi medida a pressão arterial materna e auscultado o coração do bebê, além de exames de urina e de sangue pelo menos uma vez;
- retardo de crescimento intrauterino (RCIU) - medida pela curva de crescimento de Williams (1982) - (sim, não).

As variáveis observadas categóricas foram codificadas no banco de dados de forma que a categoria de menor risco para o óbito fetal é representada pelo 0 (zero). Os níveis de exposição de maior risco são categorizados com valores maiores ou iguais a 1 (um). A tabulação dessas variáveis se encontra em anexo (Anexo 5). Essa codificação torna as relações entre os construtos e as variáveis mais facilmente interpretáveis. Para as variáveis contínuas, quanto maiores seus valores, menor o risco para o óbito fetal, exceto para densidade pessoa-cômodo que tem interpretação oposta,

3.2 Métodos

Para a análise via MEE, a partir do modelo conceitual definido, as variáveis latentes consideradas foram:

a) variável latente: situação socioeconômica (SSE)

As variáveis observadas empregadas para a construção da variável latente SSE foram:

categóricas: tipo de união (com união - codificado como 0, sem união - codificado como 1), tipo de habitação (alvenaria - 0, demais - 1), classificação socioeconômica

(A - 1, B - 2, C - 3, D - 4 ou E - 5), ocupação do chefe de família (qualificado - 0, manual - 1), plano de saúde (sim - 0, não - 1).

contínuas: renda familiar per capita (em salários-mínimos), anos de escolaridade do chefe de família, anos de escolaridade da mãe, densidade pessoa-cômodo.

b) variável latente: não-aceitação da gravidez (nAdG)

As variáveis observadas categóricas utilizadas para a construção da variável latente nAdG foram: planejamento da gravidez (sim - 0, não - 1), pensou em aborto e fez algo para abortar? (não - 0, sim - 1), reação da mãe quando soube da gravidez, do pai e da família (feliz - 0, infeliz + indiferente - 1).

Inicialmente, foi feita uma análise fatorial exploratória (AFE) em dois subconjuntos de dados gerados a partir do banco de dados de mortalidade fetal. Esse procedimento foi feito para explorar o número de fatores e sua composição. Os dois subconjuntos são selecionados aleatoriamente, sem considerar nenhuma estratificação (Hatcher 1994). Posteriormente, o modelo obtido por meio da AFE é aplicado a todo o conjunto de dados por se considerar que algumas variáveis apresentam categorias pouco frequentes na combinação entre as diversas variáveis que a metodologia utiliza.

Para os MEE, o ajuste do modelo foi obtido por meio de medidas de qui-quadrado, índice de Tucker-Lewis (TLI), índice de ajustamento comparativo (CFI), raiz do erro quadrático médio de aproximação (RMSEA), raiz do resíduo quadrático médio ponderado (Weighted Root Mean Square Residual; WRMR) (Hu e Bentler 1999).

Foram utilizados os programas Stata, versão 9.0, para a tabulação das variáveis, o que facilitou a posterior exportação de dados para o software MPlus, versão 3.01, para as análises via modelagem de equações estruturais.

As variáveis observadas união, RCIU e intercorrências foram introduzidas após o modelo de mensuração estar estabelecido.

3.2.1 Passos adotados na modelagem

Aqui estão descritos os passos adotados para fazer a modelagem.

Passo 1. Verificação do número de variáveis e composição dos fatores, via análise fatorial exploratória

Inicialmente, a proposta foi explorar o número de fatores e sua composição como uma forma de verificar a existência dos construtos pretendidos SSE e nAdG. Foi feita uma análise fatorial exploratória (AFE) em dois subconjuntos de dados gerados a partir do banco de dados de mortalidade fetal. Os dois subconjuntos foram selecionados aleatoriamente, sem considerar nenhuma estratificação (Hatcher 1994). Posteriormente, o modelo obtido por meio da AFE é aplicado a todo o conjunto de dados por se considerar que algumas variáveis apresentam categorias pouco frequentes na combinação entre as diversas variáveis que a metodologia utiliza.

Foram empregados os seguintes critérios para estabelecer o número de variáveis latentes (fatores), o mesmo número de autovalores que apresentam valores maiores do que 1 ou que a porcentagem de variância explicada acumulada seja maior do que 80% (Johnson e Wichern 1992). Após a definição do número de fatores, as cargas fatoriais são apresentadas para cada fator da AFE e valores próximos dessas cargas ajudam a definir o fator.

Para avaliar o ajuste do modelo fatorial, aplicou-se um teste para verificar se o ajuste do modelo é pior do que um modelo, em que as variáveis se correlacionam livremente. Nesse teste a distribuição qui-quadrado e p-valores acima de 5% indicam um bom ajuste (Loehlin 2004). Também foi aplicado o teste baseado no valor de RMSEA (raiz do erro quadrático médio de aproximação) para o qual, segundo Hu e Bentler (1999), os valores menores do que 0,05 são considerados como satisfatórios. Espera-se também que o procedimento não forneça uma estimativa negativa da variância dos resíduos, o que pode indicar uma má especificação do modelo (Hu e Bentler 1999).

Passo 2. Modelo com duas variáveis latentes – Situação Socioeconômica (SSE) e Não-aceitação da Gravidez (nAdG)

- o Modelo 1 do Passo 2 apresenta o mesmo modelo da análise fatorial, onde se avalia a correlação entre as variáveis latentes SSE e nAdG.

As tabelas apresentaram os interceptos das variáveis contínuas (que representam as médias dessas variáveis), os limiares (*thresholds*) para as variáveis categóricas/binárias, as variâncias das variáveis latentes e as variâncias residuais das variáveis contínuas.

Os limiares são os pontos-de-corte de cada variável contínua, com distribuição normal (com média 0 e variância 1) que, hipoteticamente, gera a variável binária observada. As variâncias das variáveis latentes foram definidas em todos os modelos com o valor 1, de modo a se obter estimativas para cada indicador. Este procedimento permite a interpretação dos coeficientes entre os indicadores e as variáveis latentes, como uma medida de correlação. A variância residual das variáveis contínuas entra no cálculo do qui-quadrado de ajuste.

- Os Modelos 2 e 3 do Passo 2, incluem possíveis correlações entre os erros das variáveis observadas baseados no efeito que cada correlação tem no índice de modificação descrito a seguir.

Foi utilizado o procedimento indicado por Schumacker e Lomax (2004), que emprega índices de modificação para representar a variação no valor da estatística qui-quadrado de ajuste com o acréscimo de cada parâmetro (correlação) no modelo. Usualmente, pode-se incluir apenas a relação que apresentar o índice de modificação mais alta (e acima de 3,84), e observar os resultados do qui-quadrado e das demais medidas de ajuste (Schumacker e Lomax 2004). Este procedimento geraram os Modelos 2 e 3. Se houver melhora do ponto de vista de ajuste do modelo, as correlações são incluídas e mantidas no modelo.

Passo 3. Modelo com as duas variáveis latentes – SSE e nAdG – e União

A variável união (com companheiro ou sem companheiro) foi introduzida no modelo a partir do Modelo 3 do Passo 2.

A união foi considerada como um fator de risco em dois estudos realizados anteriormente. No estudo da mortalidade neonatal precoce (Schoeps et al.2007), a

união foi considerada como sendo uma das variáveis que integravam o bloco psicossocial da mãe enquanto que no estudo de mortalidade fetal (Almeida et al. 2007), esta variável foi incluída no bloco de variáveis socioeconômico. Como a união pode representar tanto suporte social e econômico como suporte afetivo para a mãe durante a gestação, esta variável foi incluída no modelo.

- No Modelo 1 do Passo 3, a união foi colocada como variável explicativa de SSE e nAdG. A partir deste modelo, utilizou-se índices de modificação para definir a presença de relação entre a união (como variável resposta) e as variáveis observadas (indicadores) que foram incluídos no Modelo 2. Mesmo que as relações entre união e SSE e união e nAdG não sejam significantes, elas são mantida no modelo para obter os índices de modificação.
- No Modelo 2 do Passo 3, a união é variável explicada para os fatores SSE e nAdG, além de ser variável resposta das variáveis observadas que compõem esses fatores.
- No modelo 3 do Passo 3, foi retirada a relação entre união e SSE.

Passo 4. Modelo com SSE e nAdG, e as variáveis pré-natal, intercorrências, RCIU e o óbito fetal

- O Modelo 1 do Passo 4 inclui, a partir do Modelo 3 do Passo 3, as relações entre a ocorrência do óbito fetal e SSE, nAdG, atenção de pré-natal, intercorrências e RCIU, conforme a Figura 1.
- O Modelo Final apresenta o modelo sem as relações não significantes.

Neste trabalho, serão adotadas as seguintes medidas de ajuste do modelo (Hu e Bentler 1999) (Quadro 2).

Quadro 2. Medidas de ajuste do modelo.

medidas	valor de referência	significado dos valores das medidas	hipótese
qui-quadrado	p-valor > 0,05	-	H ₀ : modelo não ajustado
índice de Tucker-Lewis (TLI)	> 0,95	0: sem ajuste; 1: ajuste perfeito	-
índice de ajustamento comparativo (CFI)	> 0,95	0: sem ajuste; 1: ajuste perfeito	-
raiz do erro quadrático médio de aproximação (<i>Root Mean Square Error of Approximation</i> – RMSEA)	< 0,05	0: ajuste perfeito	-
raiz do resíduo médio ao quadrado ponderado (<i>Weighted Root Mean Square Residual</i> - WRMR)	< 0,90 *	0: ajuste perfeito	-

* Yu (2002) ressalta que valores próximos de 1 são mais razoáveis para respostas binárias.

4 RESULTADOS

4.1 Modelagem de equações estruturais para a mortalidade fetal

Passo 1. Verificação do número e composição dos fatores, via análise fatorial exploratória

A Figura 6 apresenta a composição inicial proposta para os fatores de situação socioeconômica e de não-aceitação da gravidez.

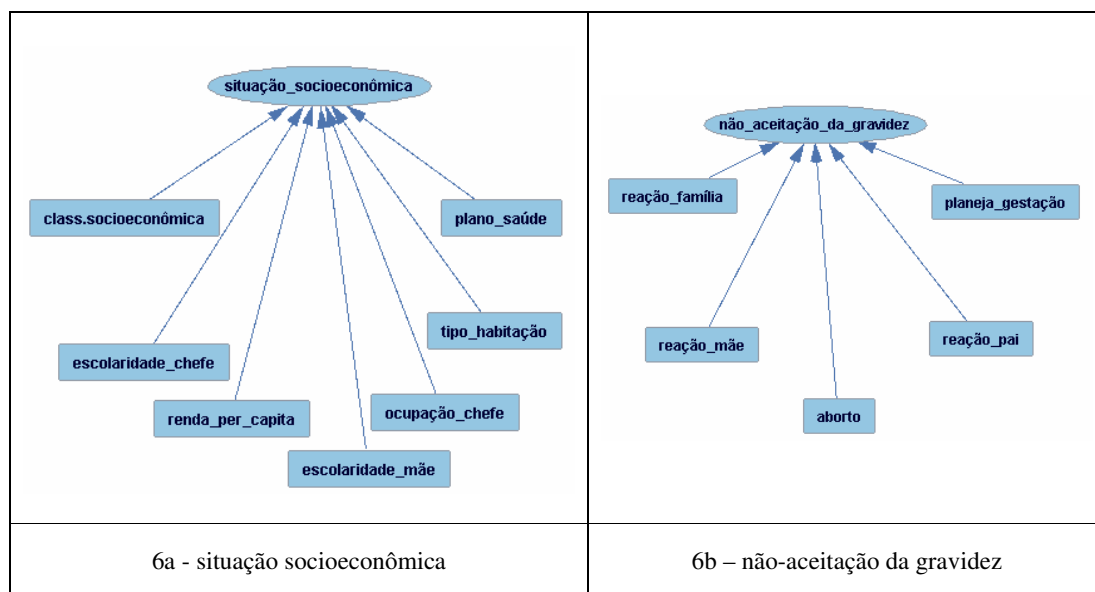


Figura 6. Modelos de mensuração das variáveis latentes.

A análise fatorial exploratória (AFE) foi aplicada para dois subconjuntos dos dados - o banco de dados foi separado em dois: o primeiro subconjunto contou com 234 indivíduos e o segundo contou com 243 indivíduos. As variáveis que entraram nesta análise foram os 13 indicadores presentes na concepção teórica das relações existentes entre as variáveis (Figuras 6a e 6b).

A Tabela 1 apresenta os resultados da AFE para esses dois subconjuntos de dados. Verifica-se que apenas 3 fatores apresentam autovalores maiores do que 1 e com porcentagem de variância explicada maior que 80% nos dois subconjuntos de dados.

Tabela 1. Autovalores da AFE para os dois subconjuntos de dados.

n. de fatores	subconjunto 1			subconjunto 2		
	autovalor	% variância explicada	% variância explicada acumulada	autovalor	% variância explicada	% variância explicada acumulada
1	4,436	53,7	53,7	5,229	61,8	61,8
2	2,918	27,5	81,2	2,554	23,4	85,2
3	1,024	0,8	82,0	1,062	1,2	86,4
4	0,859	0,06	82,6	0,817	0,1	86,5
5	0,694	0,05	83,1	0,656	0,05	86,6

A Tabela 2 apresenta os resultados dos ajustes dos modelos. Verifica-se que o teste de qui-quadrado é não significativo apenas para o modelo com 3 fatores. As outras medidas de ajuste são satisfatórias para o modelo com 3 fatores em ambos subconjuntos de dados. No modelo com 4 fatores, o valor do RMSEA é semelhante ao do valor obtido para o modelo com 3 fatores. Identificou-se, ainda, que no segundo subconjunto de dados não foi possível obter um modelo com 4 fatores.

Tabela 2. Medidas de ajuste.

n. de fatores	subconjunto 1			subconjunto 2		
	qui-quadrado	RMSEA	variância residual negativa	qui-quadrado	RMSEA	variância residual negativa
1	570,668 (65) <0,001	0,182	não	500,490 (65) <0,001	0,172	não
2	72,401 (53) 0,0395	0,040	não	85,846 (53) 0,0029	0,052	não
3	50,971 (42) 0,1615	0,030	não	49,218 (42) 0,2066	0,028	não
4	29,652 (32) 0,5859	0,040	não	- *	-	sim

RMSEA: raiz do erro quadrático médio de aproximação

* solução não foi obtida devido à presença de variância negativa dos resíduos

Tabela 3. cargas fatoriais rotacionadas (método varimax).

variáveis	subconjunto 1			subconjunto 2		
	carga do fator 1	carga do fator 2	carga do fator 3	carga do fator 1	carga do fator 2	carga do fator 3
escolaridade da mãe	0,561	0,048	0,295	0,601	-0,052	0,008
escolaridade do chefe de família	0,705	0,042	0,119	0,665	-0,12	0,015
densidade pessoa-cômodo	-0,447	0,223	-0,035	-0,580	0,132	-0,006
renda per capita	0,683	0,38	-0,170	0,712	0,231	0,012
ocupação do chefe de família	-0,575	0,047	0,049	-0,614	0,249	0,032
classificação socioeconômica	-0,665	0,02	-0,327	-0,701	0,055	0,012
plano de saúde	-0,682	0,192	-0,028	-0,849	0,378	-0,096
tipo de habitação	-0,423	-0,017	-0,727	-0,373	-0,082	0,055
reação da mãe	-0,089	0,921	-0,126	-0,132	0,734	0,203
reação da família	-0,302	0,659	0,035	-0,191	0,79	0,026
reação do pai	-0,090	0,695	0,144	-0,227	0,769	0,036
planejamento da gestação	0,058	0,878	-0,047	-0,141	0,639	0,040
tentativa de aborto	-0,122	0,754	-0,049	0,014	0,756	0,100

A Tabela 3 apresenta o resultado da composição dos 3 fatores. O conjunto composto pelas variáveis reações da mãe, pai e família, planejamento da gestação e aborto formam claramente um fator, que foi denominado “Não-aceitação da Gestação” (nAdG). O fator denominado “Situação socioeconômica” (SSE) foi composto por escolaridade da mãe e do chefe da família e renda familiar per capita. Nota-se que densidade pessoa-cômodo, ocupação do chefe de família, classificação socioeconômica, plano de saúde, tipo de habitação não parecem fazer parte de nenhum dos construtos obtidos pela AFE e também não formam um terceiro construto. Assim, optou-se pela retirada desses indicadores do modelo, tendo em vista que essas variáveis não apresentaram associação significativa com o óbito fetal em estudo anterior.

A partir dessa análise, obteve-se uma nova configuração para o modelo de mensuração, conforme representado na Figura 7, a seguir.

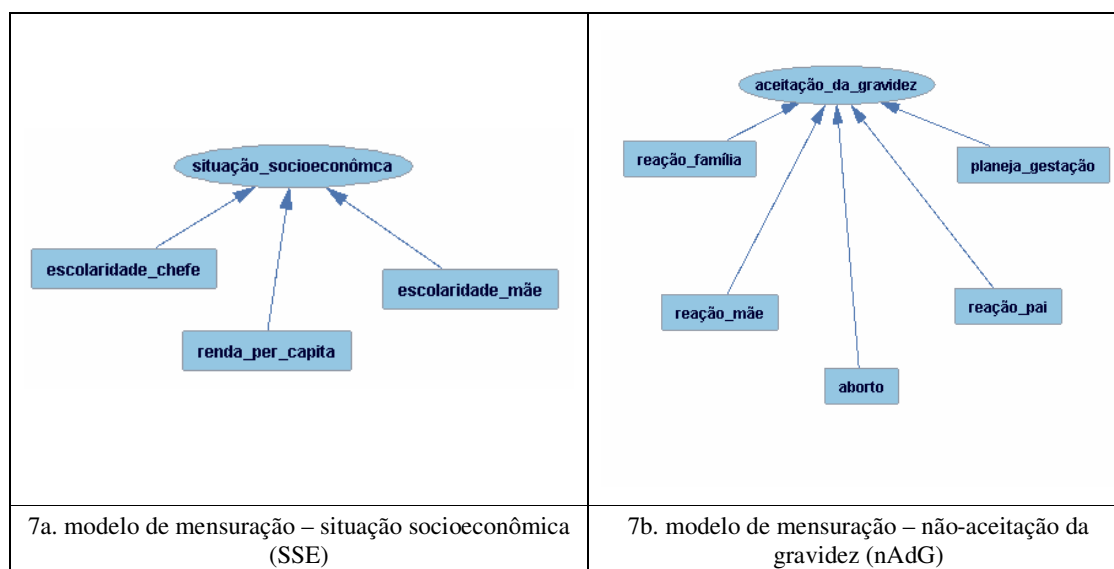


Figura 7. Diagrama de caminhos do modelo de mensuração.

Passo 2. Modelo com duas variáveis latentes – Situação Socioeconômica (SSE) e Não-aceitação da Gravidez (nAdG) –

O modelo com duas variáveis latentes foi utilizado para avaliar o quanto o modelo hipotetizado (Figura 7a) se ajusta aos dados do estudo, para tanto foi utilizada a análise fatorial confirmatória (AFC). A Tabela 4 apresenta os resultados para o modelo apresentado.

Tabela 4. Estimativas dos parâmetros para o modelo com as duas variáveis latentes SSE e nAdG – Modelo 1 do Passo 2

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
Interceptos					
escolaridade da mãe	7,359	0,155	47,414	<0,001	2,267
escolaridade do chefe de família	6,348	0,165	38,366	<0,001	1,802
renda per capita	1,978	0,066	29,908	<0,001	1,396
Limiars					
reação negativa da família	0,540	0,091	5,919	<0,001	0,540
reação negativa da mãe	1,013	0,300	3,375	<0,001	1,013
reação negativa do pai	0,857	0,103	8,334	<0,001	0,857
não-planejamento da gestação tentativa (ponto-de-corte 1)	-0,898	0,179	-5,017	<0,001	-0,898
tentativa (ponto-de-corte 2)	2,718	0,24	11,333	<0,001	2,718
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	8,049	0,820	9,810	<0,001	0,764
escolaridade do chefe de família	9,536	0,922	10,338	<0,001	0,768
renda per capita	0,112	0,356	0,315	<0,001	0,056

Com base na forma como foram categorizadas as variáveis, escores mais elevados da variável latente SSE, corresponde a valores elevados das variáveis escolaridade da mãe do chefe de família e da renda familiar per capita. Com relação à variável latente nAdG, verifica-se que escores mais elevados da não-aceitação da gravidez corresponde às reações negativas da gravidez, ao não-planejamento da gestação e à tentativa de aborto.

A correlação existente entre SSE e nAdG é pequena (-0,33), embora estatisticamente significativa. O sinal negativo indica que, em média, maiores escores de situação socioeconômica estão associados a menores escores de não aceitação da gestação. Esse resultado é razoável e indica que o aumento dos escores de SSE relaciona-se de forma oposta aos escores da nAdG.

A Tabela 5 mostra que as medidas de ajuste obtidas não estão conforme os valores de referência adotados, sugerindo que o ajuste do modelo proposto não é razoável.

Além disso, a estatística qui-quadrado também não foi satisfatória, ao apresentar p-valor bastante pequeno.

Tabela 5. Ajustamento de modelo para o modelo com SSE e nAdG – Modelo 1 do Passo 2

medidas	valores	referência
qui-quadrado (g.l.) p-valor	56,67 (14) <0,0001	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,944	>0,95
índice de Tucker-Lewis (TLI)	0,948	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,081	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	1,108	<0,90

A partir do Modelo 1 do Passo 2, calcularam-se os índices de modificação (Tabela 6). Observa-se que a inclusão da estimação das correlações existentes entre escolaridade do chefe de família e da mãe vai alterar o valor do qui-quadrado (4,698) o que influencia no ajuste do modelo, indicando a necessidade de incorporar a relação existente entre estas duas variáveis no modelo (Tabela 7).

Tabela 6. Índices de modificação para o modelo com SSE e nAdG – Modelo 1 do Passo 2.

correlações	índice de modificação
escolaridade do chefe de família com escolaridade da mãe	23,958
renda per capita com escolaridade do chefe de família	12,014
reação negativa da família com renda per capita	6,262
não-planejamento da gestação com escolaridade da mãe	5,360
não-planejamento da gestação com reação negativa da família	5,455

Tabela 7. Estimativas dos parâmetros para o modelo com SSE e nAdG – Modelo 2 do Passo 2.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
SSE					
escolaridade da mãe	1,123	0,260	4,315	<0,001	0,346
escolaridade do chefe de família	1,171	0,274	4,273	<0,001	0,332
renda per capita	1,715	0,324	5,296	<0,001	1,210
nAdG					
reação negativa da família	1,010	0,131	7,717	<0,001	0,711
reação negativa da mãe	2,502	0,731	3,422	<0,001	0,927
reação negativa do pai	1,036	0,138	7,517	<0,001	0,720
não-planejamento da gestação	1,888	0,355	5,327	<0,001	0,884
tentativa de aborto	1,078	0,157	6,87	<0,001	0,735
correlações					
SSE com nAdG	-0,311	0,074	-4,183	<0,001	-0,311
escolaridade da mãe com escolaridade do chefe de família	0,434	0,075	5,786	<0,001	0,380
Interceptos					
escolaridade da mãe	7,359	0,155	47,414	<0,001	2,267
escolaridade do chefe de família	6,348	0,165	38,366	<0,001	1,802
renda per capita	1,978	0,066	29,908	<0,001	1,396
Limiaries					
reação negativa da família	0,539	0,091	5,921	<0,001	0,539
reação negativa da mãe	1,023	0,306	3,338	<0,001	1,023
reação negativa do pai	0,856	0,102	8,354	<0,001	0,856
não-planejamento da gestação	-0,899	0,178	-5,043	<0,001	-0,899
tentativa (ponto-de-corte 1)	1,510	0,145	10,433	<0,001	1,510
tentativa (ponto-de-corte 2)	2,711	0,238	11,412	<0,001	2,711
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	9,254	0,962	9,623	<0,001	0,880
escolaridade do chefe de família	11,067	1,066	10,381	<0,001	0,890
renda per capita	0,933	1,063	0,878	<0,001	0,465

A tabela 8 mostra que as medidas de ajuste obtidas se encontram conforme os valores de referência, indicando que o ajuste do modelo proposto ainda não é razoável, porém o nível de ajuste, comparado ao anterior, é mais aceitável.

Tabela 8. Ajustamento de modelo para o modelo com SSE e nAdG – modelo 2 do Passo 2.

medidas	valores	referência
qui-quadrado (g.l.) p-valor	28,299 (13) 0,0082	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,980	>0,95
índice de Tucker-Lewis (TLI)	0,980	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,051	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,793	<0,90

A Tabela 9 apresenta índices de modificação para o Modelo 2 do Passo 2 no qual se considerou a introdução da correlação existente entre o não-planejamento da gestação e reação negativa da mãe à gestação. Esta correlação também foi incorporada ao modelo (Tabela 10).

Tabela 9. Índices de modificação para o modelo com SSE e nAdG – modelo 2 do Passo 2.

correlações	índice de modificação
não-planejamento da gestação com reação negativa da mãe	5,393
não-planejamento da gestação com reação negativa da família	5,280

Após a inclusão das correlações entre escolaridade do chefe e escolaridade da mãe, os resultados estão apresentados na Tabela 10.

Tabela 10. Estimativas dos parâmetros para o modelo com SSE e nAdG – modelo 3 do Passo 2.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
SSE					
escolaridade da mãe	1,127	0,259	4,358	<0,001	0,473
escolaridade do chefe de família	1,190	0,274	4,342	<0,001	0,437
renda per capita	1,698	0,315	5,391	<0,001	0,981
nAdG					
reação negativa da família	1,122	0,158	7,081	<0,001	0,747
reação negativa da mãe	1,601	0,307	5,218	<0,001	0,848
reação negativa do pai	1,133	0,167	6,772	<0,001	0,750
não-planejamento da gestação	1,281	0,248	5,160	<0,001	0,788
tentativa de aborto	1,169	0,189	6,186	<0,001	0,760
correlações					
SSE com nAdG	-0,325	0,076	-4,260	<0,001	-0,325
escolaridade da mãe com escolaridade do chefe de família	0,432	0,074	5,760	<0,001	0,378
planejamento da gestação com reação da mãe	0,577	0,139	4,144	<0,001	0,188

Tabela 10 (continuação). Estimativas dos parâmetros para o modelo com SSE e nAdG – modelo 3 do Passo 2.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
Interceptos					
escolaridade da mãe	7,359	0,155	47,414	<0,001	2,267
escolaridade do chefe de família	6,348	0,165	38,366	<0,001	1,802
renda per capita	1,978	0,066	29,908	<0,001	1,396
Limiões					
reação negativa da família	0,570	0,099	5,738	<0,001	0,570
reação negativa da mãe	0,716	0,146	4,923	<0,001	0,716
reação negativa do pai	0,898	0,115	7,829	<0,001	0,898
não-planejamento da gestação tentativa (ponto-de-corte 1)	-0,684	0,125	-5,456	<0,001	-0,684
tentativa (ponto-de-corte 2)	1,580	0,170	9,280	<0,001	1,580
tentativa (ponto-de-corte 2)	2,836	0,283	10,027	<0,001	2,836
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	9,254	0,960	9,637	<0,001	0,879
escolaridade do chefe de família	11,012	1,064	10,354	<0,001	0,886
renda per capita	0,875	1,020	0,858	<0,001	0,436

A tabela 11 mostra que as medidas de ajuste obtidas se encontram conforme os valores de referência adotados, indicando que o ajuste do modelo proposto é mais aceitável.

Tabela 11. Ajustamento de modelo para o modelo com SSE e nAdG – modelo 3 do Passo 2.

medidas	valores	referência
qui-quadrado (g.l.) p-valor	24,372 (13) 0,0279	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,985	>0,95
índice de Tucker-Lewis (TLI)	0,985	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,044	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,704	<0,90

Passo 3. Modelo com as duas variáveis latentes – SSE e nAdG – e União

O modelo a seguir pretende verificar a relação entre as variáveis latentes e a união e descrever as possíveis relações existentes entre as variáveis indicadoras e a união (Figura 8 e Tabela 12).

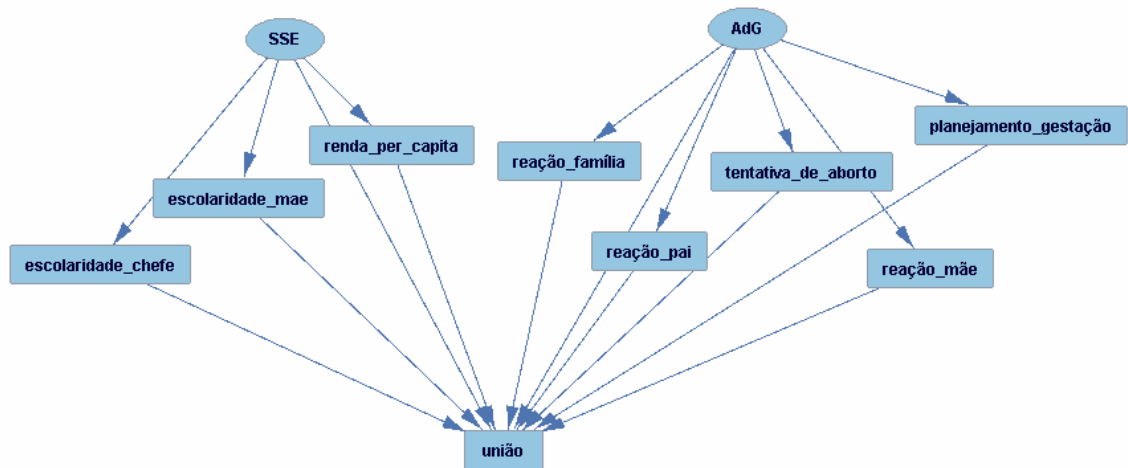


Figura 8. modelo com variável união e possíveis influências das demais variáveis.

O Modelo 1 do Passo 2 considera apenas a união em função das variáveis latentes. Neste procedimento foram avaliadas, por meio do índice de modificação, as relações existentes entre as variáveis latentes estudadas (Tabela 12). Identificou-se existir uma associação entre nAdG e união, indicando que quanto maior a não-aceitação da gestação a não união é mais presente. Não foi encontrada associação significativa entre união e SSE (Tabela 12).

Tabela 12. Estimativas dos parâmetros para o modelo com SSE, nAdG e união– Modelo 1 do Passo 3.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
correlações					
SSE com nAdG	-0,315	0,079	-4,007	<0,001	-0,315
escolaridade da mãe com escolaridade do chefe de família	0,444	0,076	5,842	<0,001	0,388
planejamento da gestação com reação da mãe	0,580	0,129	4,498	<0,001	0,197
Interceptos					
escolaridade da mãe	7,359	0,155	47,414	<0,001	2,267
escolaridade do chefe de família	6,348	0,165	38,366	<0,001	1,802
renda per capita	1,978	0,066	29,908	<0,001	1,396
Limiares					
reação negativa da família	0,586	0,103	5,718	<0,001	0,586
reação negativa da mãe	0,667	0,126	5,278	<0,001	0,667
reação negativa do pai	0,906	0,114	7,960	<0,001	0,906
não-planejamento da gestação	-0,704	0,134	-5,244	<0,001	-0,704
tentativa (ponto-de-corte 1)	1,557	0,158	9,876	<0,001	1,557
tentativa (ponto-de-corte 2)	2,794	0,259	10,771	<0,001	2,794
união	0,636	0,075	8,494	<0,001	0,636
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	9,414	0,968	9,721	<0,001	0,894
escolaridade do chefe de família	11,077	1,077	10,289	<0,001	0,891
renda per capita	1,130	1,212	0,932	<0,001	0,563

A tabela 13 mostra que a inclusão das relações entre união e SSE e nAdG não melhora o ajuste do modelo (ainda com a relação entre união e SSE não significativa), porém decidiu-se manter essas relações para avaliar as variáveis indicadoras sobre a união. A partir do índice de modificação, incluiu-se a relação da união em função de escolaridade da mãe (Tabela 14).

Tabela 13. Ajustamento de modelo para o modelo com SSE, nAdG e união – modelo 1 do Passo 3.

medidas	valores	referência
índice de ajustamento comparativo (CFI)	42,444 (17) 0,0006	p-valor>0,05
índice de Tucker-Lewis (TLI)	0,972	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,970	>0,95
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,057	<0,05
índice de ajustamento comparativo (CFI)	0,859	<0,90

Tabela 14. Índices de modificação para o modelo com SSE, nAdG e união – Modelo 1 do Passo 3.

regressões e correlações	índice de modificação
união em função de escolaridade da mãe	7,533
união em função de escolaridade do chefe de família	0,317
união em função de renda per capita	3,141
união em função de reação negativa da família	1,420
união em função de reação negativa da mãe	1,246

O modelo 2 do Passo 3 (tabela 15) inclui a relação entre união e escolaridade da mãe como variável explicativa. A inclusão feita melhora pouco o ajuste do modelo (Tabela 16) que ainda conta com a correlação não significativa entre união e SSE.

Tabela 15. Estimativas dos parâmetros para o modelo com SSE, nAdG e união – modelo 2 do Passo 3.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
SSE					
escolaridade da mãe	1,142	0,252	4,535	<0,001	0,477
escolaridade do chefe de família	1,283	0,278	4,615	<0,001	0,473
renda per capita	1,624	0,280	5,795	<0,001	0,961
nAdG					
reação negativa da família	1,181	0,167	7,070	<0,001	0,763
reação negativa da mãe	1,442	0,249	5,787	<0,001	0,822
reação negativa do pai	1,150	0,167	6,898	<0,001	0,755
não-planejamento da gestação	1,337	0,270	4,946	<0,001	0,801
tentativa de aborto	1,140	0,174	6,548	<0,001	0,752

Tabela 15 (continuação). Estimativas dos parâmetros para o modelo com SSE, nAdG e união - Modelo 2 do Passo 3.

União =					
SSE	-0,077	0,069	-1,116	0,543	-0,064
nAdG	0,621	0,100	6,235	<0,001	0,519
escolaridade da mãe	0,073	0,026	2,858	<0,001	0,200
correlações					
SSE com nAdG	-0,335	0,074	-4,553	<0,001	-0,335
escolaridade da mãe com escolaridade do chefe de família	0,394	0,074	5,324	<0,001	0,342
planejamento da gestação com reação da mãe	0,582	0,128	4,533	<0,001	0,198
Interceptos					
escolaridade da mãe	7,359	0,155	47,414	<0,001	2,267
escolaridade do chefe de família	6,348	0,165	38,366	<0,001	1,802

renda per capita	1,978	0,066	29,908	<0,001	1,396
Limiares					
reação negativa da família	0,587	0,103	5,717	<0,001	0,587
reação negativa da mãe	0,666	0,126	5,286	<0,001	0,666
reação negativa do pai	0,906	0,114	7,954	<0,001	0,906
não-planejamento da gestação tentativa (ponto-de-corte 1)	-0,702	0,134	-5,254	<0,001	-0,702
tentativa (ponto-de-corte 2)	1,557	0,158	9,860	<0,001	1,557
união	2,796	0,260	10,747	<0,001	2,796
	1,190	0,222	5,360	<0,001	1,190
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	9,402	0,968	9,711	<0,001	0,878
escolaridade do chefe de família	10,781	1,065	10,118	<0,001	0,868
renda per capita	0,631	0,854	0,738	<0,001	0,314

Tabela 16. Ajustamento de modelo para o modelo com SSE, nAdG e união – Modelo 2 do Passo 3.

medidas	valores	referência
qui-quadrado (g.l.) p-valor	35,967 (17) 0,0046	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,979	>0,95
índice de Tucker-Lewis (TLI)	0,978	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,049	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,762	<0,90

O modelo 3 do Passo 2 retira a correlação não significativa entre união e SSE (Tabela 17) e passa a apresentar medidas de ajustes razoáveis (Tabela 18).

Tabela 17. Estimativas dos parâmetros para o modelo com SSE, nAdG e união – modelo 3 do Passo 3

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
correlações					
SSE com nAdG	-0,339	0,075	-4,501	<0,001	-0,339
escolaridade da mãe com escolaridade do chefe de família	0,406	0,075	5,413	<0,001	0,351
planejamento da gestação com reação da mãe	0,585	0,126	4,636	<0,001	0,202
Limiares					
reação negativa da família	0,587	0,103	5,723	<0,001	0,587
reação negativa da mãe	0,663	0,125	5,312	<0,001	0,663
reação negativa do pai	0,905	0,114	7,970	<0,001	0,905
não-planejamento da gestação	-0,699	0,132	-5,302	<0,001	-0,699
tentativa (ponto-de-corte 1)	1,555	0,157	9,892	<0,001	1,555
tentativa (ponto-de-corte 2)	2,791	0,259	10,791	<0,001	2,791
união	1,085	0,214	5,080	<0,001	1,085

Tabela 17 (continuação). Estimativas dos parâmetros para o modelo com SSE, nAdG, união–modelo 3

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	9,519	0,976	9,756	<0,001	0,886
escolaridade do chefe de família	10,858	1,069	10,154	<0,001	0,874
renda per capita	0,723	0,911	0,794	<0,001	0,360

Tabela 18. Ajustamento de modelo para o modelo com SSE, nAdG e união – modelo 3

medidas	valores	referência
qui-quadrado (g.l.) p-valor	35,376 (17) 0,0055	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,980	>0,95
índice de Tucker-Lewis (TLI)	0,978	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,048	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,771	<0,90

Até este momento, o modelo apresentou um caráter exploratório, à medida que foram avaliados os indicadores que compõem as variáveis latentes (fatores) e a identificação do papel da variável união no modelo. A partir daqui, será testada a inclusão do pré-natal e das variáveis biológicas, tais como intercorrências médicas na gestação e retardo de crescimento intra-uterino.

Passo 4. Modelo com Situação socioeconômica e Aceitação da gravidez e pré-natal e óbito fetal

A partir do modelo anterior (Tabela 18), foram incluídas as variáveis referentes à assistência pré-natal, às variáveis biológicas (intercorrências e RCIU) e o desfecho óbito fetal. A Tabela 19 apresenta o modelo com as inclusões dessas variáveis. Observa-se que o SSE influencia a adequação da assistência pré-natal. Verificou-se que não há uma relação direta entre SSE e o desfecho, isto é, o efeito de SSE sobre os óbitos fetais manifesta-se por meio da inadequação da atenção pré-natal sugerindo que a atenção pré-natal tem efeito mediador da situação socioeconômica negativa sobre o óbito fetal. Também não se observou existir relação significativa entre a ausência de união e a inadequação do pré-natal (Tabela 19). O ajuste foi considerado razoável (Tabela 20).

Tabela 19. Estimativas dos parâmetros para o modelo com SSE, nAdG, união, pré-natal, intercorrências e RCIU sobre o óbito fetal – Modelo 1 do Passo 4

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
União =					
nAdG	0,585	0,098	5,987	<0,001	0,503
escolaridade da mãe (ESMAEC)	0,071	0,026	2,769	<0,001	0,198
Pré-natal inadequado =					
SSE	-0,322	0,089	-3,635	<0,001	-0,287
nAdG	0,260	0,106	2,447	<0,001	0,232
união	0,060	0,090	0,668	0,835	0,062
Fetal =					
SSE	-0,071	0,093	-0,758	0,755	-0,054
nAdG	0,307	0,128	2,404	<0,001	0,235
pré-natal inadequado	0,379	0,099	3,813	<0,001	0,326
intercorrências	0,288	0,102	2,810	<0,001	0,257
RCIU	0,108	0,015	7,122	<0,001	0,398
união	0,825	0,170	4,852	<0,001	0,262
correlações					
SSE com nAdG	-0,361	0,064	-5,639	<0,001	-0,361
escolaridade da mãe com escolaridade do chefe de família	0,348	0,067	5,216	<0,001	0,303
planejamento da gestação com reação da mãe	0,600	0,122	4,917	<0,001	0,214

Tabela 19 (continuação). Estimativas dos parâmetros para o modelo com SSE, nAdG, união, pré-natal, intercorrências e RCIU sobre o óbito fetal – Modelo 1 do Passo 4

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
Interceptos					
escolaridade da mãe	7,359	0,155	47,414	<0,001	2,267
escolaridade do chefe de família	6,348	0,165	38,366	<0,001	1,802
renda per capita	1,978	0,066	29,908	<0,001	1,396
Limiaries					
reação negativa da família	0,816	0,144	5,678	<0,001	0,816
reação negativa da mãe	0,723	0,155	4,676	<0,001	0,723
reação negativa do pai	0,935	0,143	6,554	<0,001	0,935
não-planejamento da gestação tentativa (ponto-de-corte 1)	-0,606	0,149	-4,078	<0,001	-0,606
tentativa (ponto-de-corte 2)	1,564	0,188	8,320	<0,001	1,564
união	2,815	0,288	9,783	<0,001	2,815
pré-natal	1,257	0,228	5,519	<0,001	1,257
óbito fetal	0,593	0,105	5,635	<0,001	0,593
óbito fetal	1,241	0,147	8,426	<0,001	1,241
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	8,742	0,933	9,368	<0,001	0,823
escolaridade do chefe de família	10,196	0,990	10,296	<0,001	0,822
renda per capita	0,085	0,386	0,221	<0,001	0,043

Tabela 20. Ajustamento de modelo para o modelo SSE, nAdG, união, pré-natal, intercorrências e RCIU sobre o óbito fetal – Modelo 1 do Passo 4

medidas	valores	referência
qui-quadrado (g.l.) p-valor	51,590 (33) 0,0207	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,981	>0,95
índice de Tucker-Lewis (TLI)	0,979	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,035	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,842	<0,90

A tabela 21 apresenta o modelo para óbito fetal excluindo a relação do SSE sobre o óbito fetal. Este modelo apresentou ajuste considerado razoável (Tabela 22).

Tabela 21. Estimativas dos parâmetros para o modelo com SSE, nAdG, união, pré-natal, intercorrências e RCIU sobre o óbito fetal – Modelo final.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
SSE					
escolaridade da mãe	1,341	0,232	5,786	<0,001	0,611
escolaridade do chefe de família	1,470	0,246	5,964	<0,001	0,417
renda per-capita	1,398	0,172	8,111	<0,001	0,908
nAdG					
reação negativa da família	1,217	0,171	7,109	<0,001	0,773
reação negativa da mãe	1,381	0,221	6,242	<0,001	0,810
reação negativa do pai	1,152	0,163	7,054	<0,001	0,755
não-planejamento da gestação	1,304	0,257	5,069	<0,001	0,794
tentativa de aborto	1,159	0,173	6,717	<0,001	0,757
União =					
nAdG	0,591	0,097	6,064	<0,001	0,506
escolaridade da mãe	0,070	0,026	2,693	<0,001	0,195
pré-natal inadequado =					
SSE	-0,320	0,086	-3,732	<0,001	-0,285
nAdG	0,301	0,086	3,516	<0,001	0,268
Fetal =					
nAdG	0,299	0,119	2,516	<0,001	0,228
pré-natal inadequado	0,434	0,104	4,177	<0,001	0,371
intercorrências	0,288	0,103	2,789	<0,001	0,256
RCIU	0,108	0,015	7,039	<0,001	0,398
união	0,830	0,171	4,841	<0,001	0,262
correlações					
SSE com nAdG	-0,358	0,063	-5,651	<0,001	-0,358
escolaridade da mãe com escolaridade do chefe de família	0,356	0,067	5,337	<0,001	0,309
planejamento da gestação com reação da mãe	0,601	0,121	4,952	<0,001	0,215
Interceptos					
escolaridade da mãe	7,358	0,204	35,987	<0,001	2,254
escolaridade do chefe de família	6,267	0,217	28,902	<0,001	1,780
renda per capita	1,990	0,087	22,900	<0,001	1,409

Tabela 21 (continuação). Estimativas dos parâmetros para o modelo com SSE, nAdG, união, pré-natal, intercorrências e RCIU sobre o óbito fetal – Modelo final.

indicadores dos fatores e variáveis	estimativa	erro-padrão (EP)	est./ EP	p	estimativa padronizada
Limiares					
reação negativa da família	0,816	0,144	5,681	<0,001	0,816
reação negativa da mãe	0,721	0,154	4,680	<0,001	0,721
reação negativa do pai	0,934	0,143	6,557	<0,001	0,934
não-planejamento da gestação tentativa (ponto-de-corte 1)	-0,606	0,148	-4,083	<0,001	-0,606
tentativa (ponto-de-corte 2)	1,563	0,188	8,332	<0,001	1,563
união	2,814	0,287	9,801	<0,001	2,814
pré-natal	1,252	0,230	5,447	<0,001	1,252
óbito fetal	0,563	0,093	6,057	<0,001	0,563
	1,236	0,147	8,408	<0,001	1,236
Variâncias					
SSE	1	0	0	1	1
nAdG	1	0	0	1	1
Variâncias residuais					
escolaridade da mãe	8,852	0,931	9,512	<0,001	0,831
escolaridade do chefe de família	10,236	0,989	10,35	<0,001	0,826
renda per capita	0,039	0,398	0,097	<0,001	0,019

Tabela 22. Ajustamento de modelo para o modelo com SSE, nAdG, união, pré-natal, intercorrências e RCIU sobre o óbito fetal – Modelo final.

medidas	valores	referência
qui-quadrado (g.l.) p-valor	51,731 (34) 0,0263	p-valor>0,05
índice de ajustamento comparativo (CFI)	0,982	>0,95
índice de Tucker-Lewis (TLI)	0,980	>0,95
raiz do erro quadrático médio de aproximação (RMSEA)	0,034	<0,05
raiz do resíduo médio ao quadrado ponderado (WRMR)	0,848	<0,90

A Figura 9 apresenta o Modelo final na forma de diagrama de caminhos.

Verifica-se que a maior carga fatorial das variáveis indicadoras do fator SSE é a renda per capita (0,97). As variáveis indicadores utilizadas na construção da variável latente nAdG apresentam cargas fatoriais semelhantes. Já se descreveu a correlação existente entre SSE e nAdG. As duas variáveis latentes influenciam a adequação do pré-natal (quanto menor o SSE, pior a adequação do pré-natal, enquanto que com o aumento da não aceitação da gravidez, observa-se aumento da inadequação).

Os resultados mostraram que a não aceitação é maior entre as mulheres sem união marital. Não há uma relação direta entre a situação socioeconômica e o estado marital materno, observa-se apenas uma relação entre a baixa escolaridade e a

ausência de união. O efeito do SSE sobre os óbitos fetais é mediado pela atenção pré-natal, enquanto que o efeito da não-aceitação da gravidez se expressa, em parte, por um efeito direto e em parte por meio do pré-natal inadequado e da ausência da união. Neste modelo, os componentes biológicos (retardo de crescimento intra-uterino e intercorrências) apresentam um efeito direto sobre os óbitos fetais (Tabela 21).

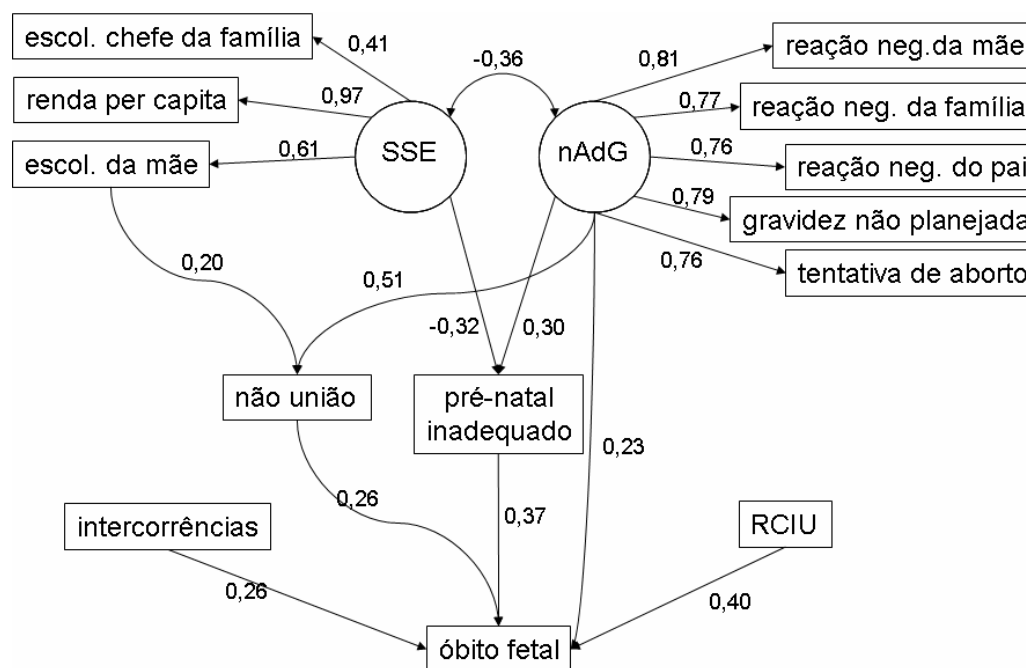


Figura 9. Modelo final para a mortalidade fetal.

De acordo com o modelo, o efeito de SSE sobre o óbito fetal ocorre via cuidado de pré-natal; o efeito de nAdG tem um componente direto sobre o óbito fetal e um componente mediado via cuidado de pré-natal. O efeito direto de intercorrências, pré-natal e RCIU representa o que foi hipotetizado e se mostrou estatisticamente significativo (Tabela 24). A tabela de efeitos mostra que os maiores efeitos foram apresentados por retardo de crescimento intra-uterino (0,40), seguido da atenção de pré-natal inadequada (0,37) e intercorrências (0,26). Ao se considerar o efeito total de nAdG, via união, pré-natal e seu próprio efeito direto, esse valor (0,47) torna esse efeito o mais influente.

Tabela 24. Efeitos diretos, indiretos e total - situação socioeconômica (SSE), não-aceitação da gravidez (nAdG), união, pré-natal inadequado, intercorrências e RCIU sobre o óbito fetal.

variável	coeficientes
<i>efeito de SSE sobre óbito fetal</i> via pré-natal	$0,37 \times -0,32 = -0,12$
<i>efeito de nAdG sobre óbito fetal</i> direto	0,23
via pré-natal	$0,30 \times 0,37 = 0,11$
via união	$0,51 \times 0,26 = 0,13$
efeito total	$0,23 + 0,30 \times 0,37 + 0,51 \times 0,26 = 0,47$
<i>efeito de intercorrências sobre óbito fetal</i> direto	0,26
<i>efeito de RCIU sobre óbito fetal</i> direto	0,40
<i>efeito de pré-natal sobre óbito fetal</i> direto	0,37
<i>efeito de união sobre óbito fetal</i> direto	0,26

4.2 Discussão

4.2.1 Discussão sobre o caminho causal para óbitos fetais

A utilização de modelos de equações estruturais engloba a construção de variáveis latentes e permite a identificação das relações entre existentes estas e as demais variáveis estudadas. O seu emprego permite identificar redes complexas de relações e, em última análise, possibilita avaliar os caminhos causais hipotetizadas pelo pesquisador. Mais do que olhar para o efeito das variáveis explicativas sobre o desfecho, busca-se olhar o caminho causal.

A utilização de variáveis latentes permite a análise de variáveis de definição complexa, como é o caso da situação socioeconômica e da aceitação da gestação, a partir de variáveis observadas. Esta metodologia permite, ainda, a identificação de variáveis mediadoras, ampliando as possibilidades de avaliar o caminho causal.

As variáveis latentes utilizadas neste trabalho são construtos hipotetizados que representam conceitos de situação socioeconômica e de aceitação da gravidez, variáveis estas geradas a partir das variáveis mensuradas no estudo sobre óbito fetal.

O modelo proposto neste trabalho mostra o papel das variáveis latentes e observadas sobre o óbito fetal.

Os MME permitem identificar o papel de variáveis complexas sobre o óbito fetal. Verificou-se que a situação socioeconômica desfavorável não influencia diretamente a ocorrência de óbitos fetais, seu efeito se dá por meio da inadequação da atenção pré-natal e da ausência de união materna.

Vários estudos anteriores haviam identificado um aumento do risco de óbitos fetais em mães e famílias com situação socioeconômica precária, avaliada por meio do efeito de variáveis, tais como baixa escolaridade (Almeida et al 2007, Stephansson 2001), ocupação manual do chefe da família (Borrel 2003), baixa renda (Aquino 1998, Menezes et al 1996). A baixa renda pode ser uma variável com ação direta sobre os desfechos em saúde, uma vez que determina as possibilidades de acesso a bens e serviços essenciais e a qualidade desses, bem como pode ser indicadora de aspectos mais complexos da posição social dos indivíduos, indicando sua posição na hierarquia de poder e propriedade na sociedade.

Os resultados desse estudo mostram que a situação socioeconômica desfavorável é melhor expressa por meio das variáveis observadas escolaridade da mãe, escolaridade da mãe, escolaridade do chefe de família e renda familiar per capita. Identificou-se, ainda, existir uma correlação entre a escolaridade do chefe da família e da mãe. Resultado semelhante foi encontrado em outros estudos (Barros et al. 2008, Schoeps et al 2007). A precária condição social do chefe de família e a baixa escolaridade materna são fatores que já haviam sido indicados como fatores de risco para o óbito fetal em alguns trabalhos (Künzel 2003, Chalumeau 2002, Stephansson 2001).

Alguns estudos também, identificaram o efeito negativo de mães vivendo sem união e da carga psicológica sobre a mãe e o óbito fetal (Künzel 2003, Chalumeau 2002, Stephansson 2001).

A variável não-aceitação da gestação foi obtida a partir das reações negativas da mãe, pai e família, não-planejamento e tentativa de aborto na gestação. Esta variável atua sobre os óbitos fetais por meio de três caminhos: um direto, um via inadequação do pré-natal e outro por meio da ausência de união materna.

A não aceitação da gestação pode ser considerada como uma variável psicossocial, pois podem gerar situações de estresse (Kramer 2003) que irão afetar a mãe, o desenvolvimento fetal e as relações sociais, principalmente as familiares.

A presença de fatores psicossociais sobre desfechos negativos da gestação já havia sido indicada por outros autores (Kramer 2003, Misra et al. 2001). No estudo de Misra et al. sobre fatores de risco para nascimentos de pré-termo, verificou-se que fatores sociais e psicossociais estavam associados ao que foi denominado fatores biomédicos (tais como doenças crônicas pré-existentes, pré-eclampsia, atenção pré-natal e hábito de fumar). Estes autores mostraram também que alguns fatores psicossociais relacionam-se com os nascimentos de pré-termo de forma independente dos fatores biomédicos. Identificou-se que a presença de estresse na gestação e a falta de controle sobre a vida estavam associadas à prematuridade. Estes autores empregaram um conjunto de medidas para avaliação da situação socioeconômica, como a escala de recursos familiares (*family resource scale*) que contemplava 30 itens que incluem tempo e dinheiro para necessidades essenciais (como moradia, transporte, cuidado com a saúde) e para necessidades não-essenciais (como viagens, tempo para si mesma). Misra et al. (2001) empregou também um questionário composto por 13 itens para avaliar o controle sobre a própria vida (*locus of control*, chamados de lócus de controle por alguns autores nacionais). Para identificar estes fatores, foi utilizada uma análise fatorial para obter os construtos socioeconômico e psicossocial. Como conclusão, esses autores sugerem que fatores psicossociais podem mediar os fatores socioeconômicos. Os resultados obtidos no presente trabalho também mostraram existir uma correlação entre as variáveis latentes SSE e nAdG, mostrando a necessidade de se levar em conta essa relação nos estudos de desfechos reprodutivos. Kramer (2003) já havia indicado que uma das causas da atenção pré-natal inadequada pode ser a presença de gestações não desejadas, que no presente estudo encontra-se representada pela não-aceitação da gravidez pela mãe, pai e família, não planejamento da gravidez e pela tentativa de aborto.

Kramer et al. (2000) resumiu fatores psicossociais, de comportamento e médicos conhecidos e associados ao baixo peso ao nascer, seja como resultado de nascimentos de pré-termo ou devido ao retardo de crescimento intra-uterino. Essas

características foram encontradas de forma mais prevalente entre os grupos com maiores limitações socioeconômicas.

Verificou-se ainda que a não-aceitação da gestação estava associada à presença de mães sem companheiro que, por sua vez, era mais freqüente entre as mães de baixa escolaridade. Outros estudos já haviam identificado a importância de mães sem companheiro na cadeia causal dos óbitos fetais (Chalumeau 2002, Künzel 2003, Stephansson 2001, Almeida et al 2007). Estes estudos mostraram também que, além de mãe sem companheiro, estão presentes condições sociais desfavoráveis tais como baixa escolaridade materna e ausência paterna (Schoeps et al. 2007, Zhang 2004, Zeitlin 2002, Kramer 2001, Stephansson 2001) e carga psicológica sobre a mãe (Kramer 2003, Misra et al. 2001).

Os resultados obtidos mostram que a atenção pré-natal apresenta um efeito direto sobre a ocorrência dos óbitos fetais e exerce papel de variável mediadora da SSE e da nAdG. Alguns estudos já haviam mostrado a associação entre ausência ou inadequação de pré-natal e os óbitos fetais (Conde-Agudelo 2000, Almeida et al. 2007, McClure 2007). Outros estudos nacionais já haviam mostrado que a baixa escolaridade era mais freqüente entre as mães com pré-natal inadequado (Leal 2004, Menezes et al. 1996). O estudo de Gama et al. (2002) mostra que a baixa instrução e renda das mulheres estavam associadas ao menor número de consultas de pré-natal, assim como fatores como hábito de fumar e consumo de drogas ilícitas. Esse estudo mostrou também que as mulheres de 20-34 anos que foram gestantes na adolescência tinham maior exposição a abortos, menor nível de escolaridade, ausência de emprego remunerado e apresentaram maior percentual de proles numerosas. Esse estudo apontou, ainda, que a assistência pré-natal se mostrou como eficiente para a prevenção da prematuridade e do baixo peso ao nascer, principalmente entre as mães adolescentes.

A presença de intercorrências, tais como presença de doença crônica materna (diabetes e problemas renais) ou de sangramento e hipertensão na gestação são fatores de risco proximais para os óbitos fetais e encontram-se longamente descritos na literatura (Chalumeau 2002, Künzel 2003, Conde-Agudelo 2000, Zhang 2004). Observou-se que estas condições apresentavam efeito direto sem ser mediado pelo

cuidado pré-natal. Esse resultado, no entanto, pode ser motivado pelo desenho do estudo, que é do tipo caso-controle. Nestes estudos não fica claro, por falta de informações, a seqüência de ocorrência das intercorrências. Este fato dificulta a avaliação se a atenção pré-natal pode minimizar os efeitos da hipertensão ou de problemas que podem levar à presença de sangramento.

Os trabalhos de Salihu et al. (2006), Zhang (2004) e Zeitlin (2002) mostram que a presença de RCIU é fator de risco para o óbito fetal, além de ser um marcador para óbitos fetais posteriores/subseqüentes. Além disso, Salihu et al. (2006) mostram que o risco aumenta com o decréscimo da idade gestacional dos nascimentos com retardo de crescimento intra-uterino (RCIU). Além de apresentar o RCIU como fator de risco, este trabalho também sugere o acompanhamento de casos com características de gestações de baixo peso anterior a partir do pré-natal.

No presente trabalho, ao apresentar as hipóteses de que havia efeito direto da situação socioeconômica e da aceitação da gravidez sobre os óbitos fetais, e também que havia um caminho via atenção pré-natal. O efeito desses dois construtos sobre o óbito fetal foi significativo via a variável de mediação, o cuidado de pré-natal, além do efeito direto de aceitação da gravidez, em um modelo que contou com a presença de intercorrências (complicações médicas) e RCIU. Esse resultado pode sugerir um papel bastante importante ao pré-natal, principalmente, para mães em situações menos favorecidas econômica e socialmente.

4.2.2 Discussão sobre a metodologia utilizada

O quadro conceitual utilizado em modelos de regressão logística contém uma lógica de relação do efeito dos blocos sobre o desfecho e relações entre os blocos existentes entre as variáveis dentro de cada bloco e pertencentes a blocos distintos. Contudo, não é possível detalhar as relações existentes entre as variáveis. Estudos com diversas variáveis (e poucos indivíduos) instigam a busca de metodologias mais flexíveis, que permitem acessar o efeito de fatores mais distais e, principalmente, acessar os efeitos de dimensões socioeconômicas sobre o desfecho. Nos modelos de regressão logística usuais admitem apenas um desfecho. Não sendo possível identificar desfechos intermediários ao longo do caminho causal.

A construção de um quadro conceitual, feita desde o início da pesquisa, auxilia a entender quais informações se quer obter, ajudando, assim, a pensar na forma de obtenção dos dados a ser utilizada na pesquisa e na análise dos dados. Deve-se notar, no entanto, que algumas relações no quadro conceitual inicial podem ser alteradas à medida que as primeiras análises são realizadas e a partir dos resultados obtidos em decorrência dos testes dos modelos propostos. Isso é especialmente provável (e até desejável) nas situações em que não se tem ainda conhecimento vasto na literatura a respeito do problema estudado e de características na população estudada, que podem diferir de outras já encontradas na literatura.

Quando uma variável de confusão é antecedente no tempo à variável de exposição, então a variável de exposição (que pode ser considerada como causal) passa a ser uma variável intermediária e a variável de confusão passa a ser considerada a variável causal (Susser 1973). Sempre se pressupõe que a variável explicativa é antecedente no tempo à variável inicialmente considerada como causal. Alguns autores (Scklo 2000, Gordis 2004) consideram essa situação de forma diferente. Se a variável de confusão for antecedente no tempo à variável de exposição, então a variável de exposição é considerada como sendo uma variável de mediação. Sob esta ótica de antecedência, o caminho proposto pelo quadro conceitual pode não ser o único, e as modelagens podem gerar resultados e interpretações diferentes.

A análise exploratória é mais importante quando o quadro conceitual não é suficientemente claro ou conhecido justificando a exploração empírica de diferentes caminhos, como é o caso dos óbitos fetais.

A proposta de testar partes do modelo (p.ex., calculando a análise fatorial confirmatória para cada variável latente (modelo de mensuração) e posteriormente testar o modelo com as variáveis latentes (modelo estrutural), não é bem aceita por alguns autores (Mulaik 2000, Hayduk et al. 2000), devido a não existência de testes adequados para analisá-los separadamente. Segundo Hayduk et al. (2000), deveria ser testado o modelo final completo, contemplando os modelos de mensuração e modelo estrutural. No presente estudo sobre a mortalidade fetal, a primeira etapa realizada a partir da análise fatorial foi importante para validar as variáveis latentes.

As variáveis latentes obtidas não haviam sido propostas inicialmente no projeto de pesquisa, ou seja, antes da fase de coleta de dados e esta necessidade surgiu com a realização de estudos anteriores e outros estudos mais recentes que levantaram novas perguntas e novas possibilidades de estudo. Esta necessidade surgiu a partir das dúvidas geradas em análises anteriores sobre o papel de algumas variáveis, mais especificamente, se era possível considerar a assistência pré-natal como variável mediadora. Apenas estudos epidemiológicos mais recentes abordavam claramente a questão das variáveis mediadoras e o emprego de MME como uma possibilidade de testar quadros conceituais.

A utilização de subgrupos de análise (como feito aqui) para o teste de modelos propostos pode auxiliar na verificação da validade de modelos de equações estruturais. É necessário, contudo, prestar atenção para o número de fatores envolvidos no quadro conceitual, pois, além do tamanho da amostra necessária, ele deve ser razoavelmente bem fundamentado para garantir que o modelo seja testado em outros estudos.

Neste trabalho, os testes realizados na fase exploratória indicaram a existência de três fatores nos dois subconjuntos de indivíduos considerados. Observou-se, contudo, que apenas duas variáveis latentes foram admitidas, SSE e nAdG, que são aquelas que têm uma interpretação mais clara. O terceiro fator não parece representar nenhum construto. Além disso, alguns epidemiologistas têm criticado a inclusão de variáveis que não possuem uma explicação clara nos modelos (Greenland et al 1986). Erros-padrão e limites de confiança são pouco apresentados em MEE e a apresentação desses valores podem dar uma idéia, por exemplo, da falta de variáveis relevantes ou da especificação correta do modelo.

Os modelos de regressão são interessantes para identificar variáveis explicativas, que podem ser preditoras para o evento estudado, e desse modo, auxiliam a identificação de indivíduos que apresentam maior risco ou efeitos de grupos. Uma das características dos MEE é o estudo de sistemas, em que há a tentativa de inclusão de múltiplos processos para o desfecho estudado, e a possibilidade de visualizar como estes se relacionam entre si. Esse processo pode indicar respostas para as questões

epidemiológicas referentes às relações de associação e causa que podem existir em diversos níveis de exposição.

Os modelos de equações estruturais auxiliam a organizar e a estimar relações que, usualmente são realizadas de forma separada. De certa forma, essa metodologia obriga a testar as hipóteses levantadas e, para tanto, é necessário se dispor de um conjunto de dados razoáveis para realizar esta tarefa. A experiência deste trabalho sugere a necessidade de se aprimorar os instrumentos de medida que devem ser planejados antes de se levar a campo o estudo.

Não estão disponíveis muitas ferramentas de análise de diagnóstico para modelos de equações estruturais, o que poderia ser particularmente interessante para a detecção de presença de subpopulações no estudo – essas subpopulações poderiam apresentar composição de fatores e efeitos diferentes sobre o desfecho estudado, o que geraria modelos distintos. Estudos nesse sentido devem ser feitos para se identificar possíveis subpopulações.

4.3 Conclusões

Com a metodologia proposta pelo MEE foi possível:

- 1) identificar a presença de construtos (variáveis latentes). Os construtos encontrados foram denominados situação socioeconômica e não-aceitação da gestação, no caminho causal dos óbitos fetais. Foi possível ainda identificar que a não-aceitação da gestação age por três mecanismos distintos e a situação socioeconômica é mediada pela atenção de pré-natal sobre o óbito fetal;
- 2) verificar que a inadequação pré-natal apresenta efeito direto sobre os óbitos fetais e também desempenham o papel de variável mediadora da situação socioeconômica e não-aceitação da gestação;
- 3) identificar um efeito direto de intercorrências e retardo de crescimento intra-uterino sobre o óbito fetal;
- 4) verificar que a união expressa, em parte, a escolaridade materna e a não-aceitação da gestação.

5 REFERÊNCIAS

- ABEP - Associação Brasileira de Empresas de Pesquisa. *Critério de Classificação Econômica Brasil*. 2000. Disponível em <<http://www.abep.org/?usaritem=arquivos&iditem=23>>. Acessado em [12 dez 2008].
- Aquino TA, Sarinho SW, Guimarães MJB. Fatores de risco para a mortalidade perinatal no Recife - 2003. *Epidemiol. Serv. Saúde* 2007;16(2):132-135. ISSN 1679-4974. Disponível em <http://scielo.iec.pa.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742007000200012&lng=pt&nrm=iso>.
- Almeida MF, Alencar GP, Novaes HMD, Siqueira AAF, Campbell O, Schoeps D, Rodrigues LC. Risk-factors for antepartum fetal deaths in the city of São Paulo, Brazil. *Revista de Saúde Pública / Journal of Public Health*, v. 41, p. 35-43, 2007.
- Almeida MF, Alencar GP, Novaes HMD, Ortiz LP. Sistemas de informação e mortalidade perinatal: conceitos e condições de uso em estudos epidemiológicos. *Revista de Saúde Pública* 2006; 9(1):56-68.
- Barata RB. Epidemiologia social. *Revista Brasileira de Epidemiologia*. 2005; 8(1):7-17.
- Brook W, Finch C. Aggression, intrapsychic distress and drug use: antecedent and intervening processes. *Journal of the American Academy of Child and Adolescent Psychiatry* 1995; 34(8):1076-1083.
- Borgoni R, Berrington AM, Smith WF. *Selecting and fitting graphical chain models to longitudinal data*. Southampton, UK, Southampton Statistical Sciences Research Institute, 33pp. (S3RI Methodology Working Papers, M04/05).

2004. Disponível em <<http://eprints.soton.ac.uk/8178/>> Acessado em [12 maio 2008].
- Borrell C, Cirera E, Ricart M, Pasarim MI, Salvador J. Social inequalities in perinatal mortality in Southern European city. *Eur J Epidemiol.* 2003; 18:5-13.
- Bollen KA. *Structural equations with latent variables.* New York: Wiley. 1989.
- Bollen KA, Lennox R. Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin* 1991; 110:305-14.
- Brown TA. *Confirmatory Factor Analysis for Applied Research.* Guilford Press, 2006. 475p.
- Browne MW. Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal* 1974; 8:1- 24.
- Browne MW. Covariance structures. In Hawkins DM. (Ed.) *Topics in applied multivariate analysis* (pp. 72-141). Cambridge: Cambridge University Press. 1982.
- Bussab WO, Morettin PA. *Estatística Básica.* Ed. Saraiva: São Paulo. 2002.
- Barros MBA; César CLG; Carandina L, Goldbaum M. *As dimensões da Saúde - Inquérito populacional em Campinas.* 1 ed. São Paulo: HUCITEC. 2008. v. 1. 229 p.
- Clayton D, Hills M. *Statistical Model in Epidemiology.* New York: Oxford University Press. 1993.
- Dunteman GH. *Principal Components Analysis.* Newbury Park: Sage Publications; 1989.
- Finney S, DiStefano C. Nonnormal and categorical data in structural equation modeling. In: Hancock GR, Mueller RO. *Structural equation modeling: a*

second course. Quantitative methods in education and the behavioral sciences. Connecticut: IAP. 2006.

Forssas E, Gissler M, Sihvonen M, Hemminki E. Maternal predictors of perinatal mortality: the role of birthweight. *International Journal of Epidemiology* 1999; 28: 475-478.

Fuchs SC, Victora CG, Fachel J. Modelo hierarquizado: uma proposta de modelagem aplicada à investigação de fatores de risco para diarreia grave. *Revista de Saúde Pública* 1996; 30(2):168-78.

Freedman D. From association to causation via regression. *Advances in Applied Mathematics* 1997; 18:59-110.

Gama SGN, Szwarcwald CL, Leal MC. Experiência de gravidez na adolescência, fatores associados e resultados perinatais entre puérperas de baixa renda. *Cad. Saúde Pública* 2002; 18(1):153-161.

Goldenberg R, Culhane JF. Low birth weight in the United States. *The American Journal of Clinical Nutrition* 2007; 85(supp):84S-90S.

Gordis L. *Epidemiology*. Philadelphia: W.B.Saunders. 2 ed. 2004.

Greenland S. Modelling and variable selection in epidemiologic analysis. *American Journal of Public Health* 1989; 79(3):340-349.

Greenland S, Brumback B. An overview of relations among causal modelling methods. *International Journal of Epidemiology* 2002; 31:1030-1037.

Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999; 10:37-48.

- Greenland S, Schlesselman JJ, Criqui MH: The fallacy of employing standardized regression coefficients and correlations as measures of effect. *Am J Epidemiol* 1986, 2:203-208.
- Hair J, Black B, Babin B, Anderson RE, Tatham RL. *Multivariate Data Analysis*. 5ed. New York: Prentice Hall. 2005.
- Hamilton JD. *Time Series Analysis*. Princeton University Press. 820 p. 1994.
- Harrell FE. *Regression modeling strategies*. New York: Springer. 2001.
- Hatcher L. *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute Press. 1994.
- Hayduk LA, Glaser DN. Doing the Four-Step Right, Right 2-3, Wrong 2-3: a brief reply to Mulaik and Millsap; Bollen; Bentler; and Herting and Costner. *Structural Equation Modeling: a multidisciplinary journal* 2000; 7(1):111-123.
- Hill AB. *Principles of Medical Statistics*. 9 ed. Oxford University Press 1971.
- Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York: Wiley-Interscience Publication. 1989.
- Hox JJ, Bechger TM. An Introduction to Structural Equation Modeling. *Family Science Review* 1998; 11:354-373.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 1999; 6:1-55.
- James LR, Mulaik SA, Brett JM. *Causal Analysis*. Beverly Hills: Sage Publications. 1982.

- Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. 3ed. New Jersey: Prentice Hall. 1992. 642p.
- Kaplan D. *Structural Equation Modeling: foundations and extensions*. Sage: Thousand Oaks. 2000.
- Kim K, Timm N. *Univariate and multivariate general linear models: theory and applications with SAS*. 2.ed. Chapman & Hall/CRC: New York. 2007.
- Klem L. Path analysis. In: Grimm LG, Yarnold PR. *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association, 1995.
- Kramer MS, Séguin L, Lydon J, Goulet L. Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? *Paediatr Perinat Epidemiol*. 2001; 14(3):194-210.
- Kramer MS. The Epidemiology of Adverse Pregnancy Outcomes: An Overview. *The Journal of Nutrition* 2003; 133:S1592-S1596.
- Koifman S, Koifman RJ. Incidência de câncer de estômago no Brasil: estudo ecológico com fatores de risco selecionados. *Cad. Saúde Pública* 1997;13(supl.1):85-92. Disponível em <<http://www.scielo.br/pdf/csp/v13s1/1428.pdf>>
- Leal MC, Gama SGN, Ratto KMN, Cunha CB. Uso do índice de Kotelchuck modificado na avaliação da assistência pré-natal e sua relação com as características maternas e o peso do recém-nascido no município do Rio de Janeiro. *Cad Saúde Pública*. 2004; 20:63-72.
- Loehlin JC. *Latent variables models: an introduction to factor, path and structural analysis*. 3. ed. New Jersey: Lawrence Erlbaum, 1998.

- Luiz RR, Struchiner CJ. *Inferência causal em epidemiologia: o modelo de respostas potenciais*. Rio de Janeiro: Editora Fiocruz. 2002.
- Macmahon B, Pugh T. *Epidemiology - principles and methods*. 1ed. Little, Brown and Co.: Boston, 1970.
- Magalhães MN, Lima ACP. *Noções de probabilidade e estatística*. 4. ed. São Paulo: EDUSP. 2002.
- McCullagh P, Nelder JA. *Generalized Linear Models*. 2 ed. Chapman and Hall, London. 1989.
- McClure EM, Wright LL, Goldenberg RL, Goudar SS, Parida SN, Jehan I et al. The Global Network: A Prospective Study of Stillbirths in Developing Countries. *Am J Obstet Gynecol*. 2007; 197(3): 247.e1–247.e5.
- Menezes AMB et al. Mortalidade perinatal em duas coortes de base populacional no Sul do Brasil: tendências e diferenciais. *Cad. Saúde Pública* 1996; 12(suppl.1).
- Misra DP, O’Campo P, Strobino P. Testing a sociomedical model for preterm delivery. *Paediatric and Perinatal Epidemiology* 2001; 15: 110-122.
- Muthén LK, Muthén BO. How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Structural Equation Modeling: a multidisciplinary journal* 2002;9(4):599-620.
- Muthén B. A general structural model with dichotomous, ordered categorical and continuous latent variable indicators. *Psychometrika* 1984; 49:115–132.
- Muthén LK, Muthén BO. *Mplus User’s Guide*. Third Edition. Los Angeles, CA: Muthén & Muthén (2004).

- Mulaik SA, Millsap RE. Doing the Four-Step Right. *Structural Equation Modeling: a multidisciplinary journal* 2000; 7(1):36-73.
- Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied Linear Statistical Analysis*. 4. ed. Chicago: Irwin. 1408p. 1996.
- Olinto MTA. Reflexões sobre o uso do conceito de gênero e/ou sexo na epidemiologia: um exemplo nos modelos hierarquizados de análise. *Revista Brasileira de Epidemiologia* 1998; 1(2):161-169.
- Organização Mundial da Saúde (World Health Organization). *Neonatal and perinatal mortality: Country, regional and global estimates*. World Health Organization: Geneva: 2006.
- Pearl J. *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press. 2000.
- Pereira LAA, Loomis Dana, Conceição GMS, Braga ALF, Arcas RM, Kishi HS et al. Association between Air Pollution and Intrauterine Mortality in São Paulo, Brazil. *Environmental Health Perspectives* 1998; 106(6).
- Reime B, Ratner PA, Tomaselli-Reime SN, Kelly A, Schuecking BA, Wenzlaff P. The role of mediating factors in the association between social deprivation and low birth weight in Germany. *Social Science & Medicine* 2006; 62:1731–1744.
- Rothman KJ, Greenland S. *Modern epidemiology*. 2 ed. Philadelphia: Lippincott-Raven. 1998.
- Rothman KJ. *Modern epidemiology*. Boston: Little Brown, 1986.

- Rubin D. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology* 1974; 66(5): 688-701.
- Salihu HM, Sharma PP, Alyiu MH, Kristensen S, Grimes-Dennis J, Kirby RS, Smulian J. Is small for gestational age a marker of future fetal survival in utero? *Obstetrics and Gynecology* 2006; 107(4):851-856.
- Schoeps D, Almeida MF, Alencar GP, França Junior I, Novaes HMD, Siqueira AAF, Campbell O, Rodrigues LC. Risk factors for early neonatal mortality. *Revista de Saúde Pública / Journal of Public Health*, v. 41, p. 1, 2007.
- Skrondal A, Rabe-Hesketh S. *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC. 2004.
- Susser M. *Causal thinking in the health sciences concepts and strategies of epidemiology*. New York: Oxford University Press. 1973.
- Szklo M, Nieto FJ. *Epidemiology: beyond the basics*. Aspen. 2000.
- Victora CG, Huttly SR, Fuchs SC, Olinto MTA. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. *International Journal of Epidemiology* 1997; 26(1):224-227.
- Zeitlin JA, Saurel-Cubizolles MJ, Ancel PY, EUROPOP Group. Marital status, cohabitation, and risk of preterm birth in Europe: where births outside marriage are common and uncommon. *Paediatr Perinat Epidemiol*. 2002; 16(2):124-30.
- Zhang J, Klebanoff MA. Small-for-gestational-age infants and risk of fetal death in subsequent pregnancies. *N Engl J Med*. 2004; 350(8):754-6.

Williams RL, Creasy RK, Cunningham GC, Hawes WE, Norris FD, Tashiro M. Fetal growth and perinatal viability in California. *Obstetrics & Gynecology* 1982; 59: 624-32.

Wooldridge JM. *Introductory econometrics: a modern approach*. 6ed. Ohio: Thomson. 2003.

Yu C-Y, Muthén B. *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes*. Technical report. 2002. Disponível em [<http://www.statmodel.com/download/Yudissertation.pdf>] Acessado em [18 ago 2007].

Yu CH. Causation in quantitative research methodologies from path modeling, SEM to TETRAD. *Theory and Science* 2007: 9(20). Disponível em <<http://theoryandscience.icaap.org/content/vol9.3/chong.html>>. Acessado em [18 ago 2007].

6 QUESTÕES ÉTICAS

O desenvolvimento do estudo segue os requisitos da Resolução 196/96 do Conselho Nacional de Saúde/Ministério da Saúde do Brasil (Comissão Nacional de Ética em Pesquisa 2000) e as normas do Comitê de Ética da Faculdade de Saúde Pública da Universidade de São Paulo, que regulamentam pesquisas envolvendo seres humanos. O projeto de pesquisa matriz foi previamente aprovado pelo Comitê de Ética da Faculdade de Saúde Pública da Universidade de São Paulo, em 15 de setembro de 2001 (Anexo 6).

ANEXO 1 – ESCRIVENDO O MODELO DE EQUAÇÕES ESTRUTURAIS

O modelo de mensuração é representado por um sistema de duas equações, uma para as variáveis endógenas e outra para as variáveis exógenas. As seguintes equações descrevem a relação de Y e X , em que Y representa as variáveis endógenas observadas e X representa variáveis exógenas observadas e são escritas da forma matricial que segue:

$$[1] \quad y = \Lambda_y \eta + \varepsilon, \text{ modelo de mensuração de } Y,$$

$$[2] \quad x = \Lambda_x \xi + \delta, \text{ modelo de mensuração de } X,$$

$$[3] \quad \eta = B\eta + \Gamma\xi + \zeta, \text{ modelo estrutural,}$$

em que η (eta) é o vetor das variáveis latentes dependentes; ξ (csi) é o vetor das variáveis latentes independentes; B (beta) é a matriz de coeficientes, cujos elementos representam os efeitos diretos das variáveis η em outras variáveis η ; Γ (gama) é a matriz de coeficientes, cujos elementos representam os efeitos diretos das variáveis ξ em variáveis η ; ζ (zeta) é o vetor aleatório de resíduos (erro na equação estrutural); y e x são os vetores de variáveis observadas; Λ_x e Λ_y são as matrizes de coeficientes, cujos elementos representam o efeito das variáveis η nas variáveis y e das variáveis ξ nas variáveis x ; ε e δ são vetores de erro (erro de mensuração) de y e x , respectivamente.

Além dessas equações, há também duas matrizes de covariâncias associadas aos resíduos das variáveis do modelo de mensuração, que representaremos por Θ_δ e Θ_ε . $\Theta_\delta = \text{Var}(\delta)$ é uma matriz $q \times q$, e $\Theta_\varepsilon = \text{Var}(\varepsilon)$ é uma matriz $p \times p$. Considerando-se a análise fatorial, essas matrizes seriam diagonais contendo a variância associadas às variáveis. Algumas suposições são aplicadas: $E[\varepsilon]=0$ e $E[\delta]=0$; ε é não correlacionado com η ($\text{cov}(\varepsilon, \eta)=0$); δ é não correlacionado com ξ

($\text{cov}(\delta, \xi)=0$); ε é não correlacionado com ξ ($\text{cov}(\varepsilon, \xi)=0$); ε é não correlacionado com δ ($\text{cov}(\varepsilon, \delta)=0$); δ é não correlacionado com η ($\text{cov}(\delta, \eta)=0$).

Para o modelo estrutural, as matrizes de covariâncias são: Φ (fi) é a matriz de covariâncias dos fatores exógenos ξ , e Ψ (psi) é a matriz de covariâncias dos resíduos das variáveis latentes ζ . Aqui, assume-se que $E[\xi]=0$, $E[\zeta]=0$ e $E[\eta]=0$, além dos resíduos das variáveis latentes ζ (chamado também de *distúrbio*; *disturbance*) e das variáveis latentes ξ serem não correlacionadas ($\text{cov}(\xi, \zeta)=0$) e que $I-B$ é não singular, $I-B$ admite inversa e I é matriz identidade $m \times m$. A tabela a seguir apresenta os símbolos das variáveis mencionadas aqui.

Pode-se ainda incluir interceptos em todas as equações e, assim, tem-se:

$$[4] y = v_y + \Lambda_y \eta + \varepsilon, [5] x = v_x + \Lambda_x \xi + \delta, [6] \eta = \alpha + B\eta + \Gamma\xi + \zeta.$$

O quadro 1 apresenta os itens a serem utilizados nos modelos a seguir.

Quadro 1. Símbolos usados nos gráficos.

itens	letras	símbolos
variáveis observadas	x, y	retângulo
variáveis latentes	η, ξ	círculo
erros	$\zeta, \varepsilon, \delta$	sem símbolo
coeficientes de regressão	$B, \Gamma, \Lambda_x, \Lambda_y$	setas com uma ponta
covariâncias	$\Phi, \Psi, \Theta_\delta$ e Θ_ε	setas com duas pontas

Regressão linear

O modelo de regressão linear múltipla pode ser representado no seguinte diagrama (Figura A1.1). Para evitar a poluição do gráfico, os ovais dos erros das variáveis serão omitidos nos exemplos apresentados.

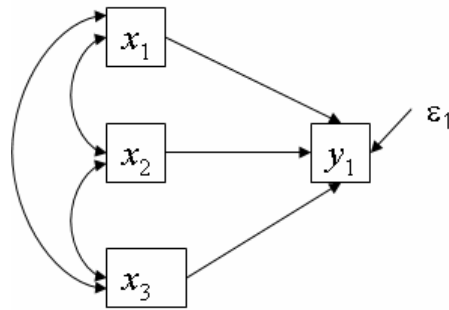


Figura A1.1. Exemplo de regressão múltipla.

Para três variáveis explicativas x_1 , x_2 e x_3 , o modelo para y_1 pode ser escrito na forma

matricial como $(y_1) = (\gamma_{11} \quad \gamma_{12} \quad \gamma_{13}) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + (\varepsilon_1)$, em ε_1 é o erro associado a y_1 .

A matriz de covariâncias de x é escrita como $\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix}$ e a variância do

erro é escrito como: $\Theta_\varepsilon = (\theta_{\varepsilon_1})$.

Análise de caminhos

A análise de caminhos não possui variáveis latentes e um exemplo com 2 variáveis explicativas y_1 e y_2 e três variáveis explicativas x_1 , x_2 e x_3 é apresentado (Figura A1.2). A Figura A1.2 representa a influência das variáveis x sobre y_2 , mediado por y_1 (efeito indireto de x sobre y_2). Aqui, nenhuma variável explicativa x não tem efeito direto sobre y_2 .

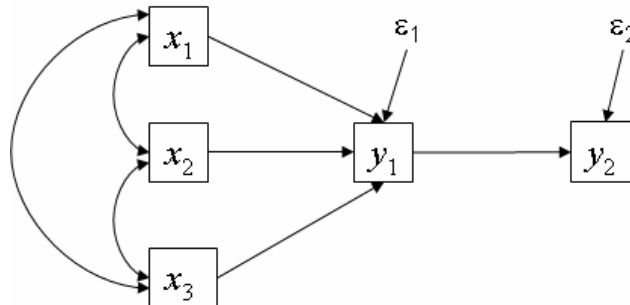


Figura A1.2. Exemplo de análise de caminhos.

Na forma matricial, temos que
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ b_{21} & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix},$$

em que ϵ_1 e ϵ_2 são erros de y_1 e y_2 , respectivamente, e b_{21} é o coeficiente de

regressão de y_2 em y_1 . A matriz de covariâncias de x é: $\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix}$ e a

matriz de covariâncias do erro é escrito como $\Theta_\epsilon = \begin{pmatrix} \theta_{\epsilon_{11}} & 0 \\ 0 & \theta_{\epsilon_{22}} \end{pmatrix}$.

Note que se o primeiro termo após o sinal de igual for retirado do modelo na forma matricial, teremos o modelo de regressão apresentado anteriormente.

Análise fatorial

Um exemplo de representação que pode ser escrito no MEE é a análise fatorial confirmatória (AFC) que é representada na Figura A1.3. Basicamente, isso corresponde ao modelo da seguinte forma:

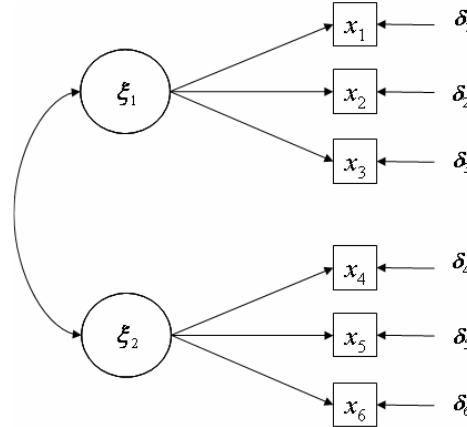


Figura A1.3. Exemplo de análise fatorial confirmatória.

No exemplo, $x = \Lambda_x \xi + \delta$ corresponde a

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} \lambda_{x_{11}} & 0 \\ \lambda_{x_{21}} & 0 \\ \lambda_{x_{31}} & 0 \\ 0 & \lambda_{x_{42}} \\ 0 & \lambda_{x_{52}} \\ 0 & \lambda_{x_{62}} \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix}.$$

A matriz de covariâncias do erro para x é escrita como

$$\Theta_\delta = \begin{pmatrix} \theta_{\delta_{11}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_{\delta_{22}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{\delta_{33}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{\delta_{44}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{\delta_{55}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{\delta_{66}} \end{pmatrix}.$$

A matriz de covariâncias das duas variáveis latentes, ξ_1 e ξ_2 , é $\Phi = \begin{pmatrix} 1 & \phi_{12} \\ \phi_{21} & 1 \end{pmatrix}$. É

necessário que as variâncias $\phi_{11} = \phi_{22} = 1$ sejam fixadas para satisfazer o critério de identificação do modelo (Anexo 2).

Modelo de equações estruturais com variáveis latentes

Para exemplificar o modelo, tomou-se o modelo de Jöreskog-Keesing-Wiley (Kim e Timm 2007) (Figura A1.4), com 3 variáveis latentes (η_1, η_2 e ξ_1), cada qual com três indicadores exógenos x_1, x_2, x_3 para η_1 ; três indicadores endógenos y_1, y_2, y_3 para η_2 e três indicadores endógenos y_4, y_5, y_6 para ξ_1 . É possível chamar as variáveis x_1, x_2, x_3 e y_1, \dots, y_6 de variáveis explicativas (e usar apenas uma letra, p.ex. y), porém diversos

livros-texto diferenciam a variável como exógena ou endógena, então optou-se por usar estes conceitos.

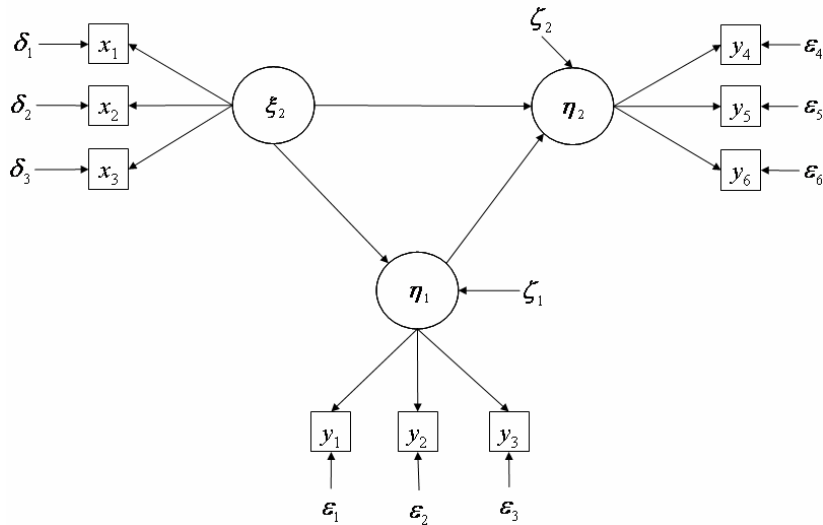


Figura A1.4. Exemplo de diagrama de caminhos para o MEE.

O MEE tem duas partes:

- O **modelo de mensuração** de Y é, matricialmente, $y = \Lambda_y \eta + \varepsilon$, e pode ser escrito como

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \gamma_{y_{21}} & 0 \\ \gamma_{y_{31}} & 0 \\ 0 & 1 \\ 0 & \gamma_{y_{52}} \\ 0 & \gamma_{y_{62}} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix},$$

e o modelo de mensuração de X é, matricialmente, $x = \Lambda_x \xi + \delta$, e pode ser escrito por

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_{x_{21}} \\ \lambda_{x_{23}} \end{pmatrix} (\xi_1) + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix}.$$

Aqui, as cargas de x_1 em ξ_1 , y_1 em η_1 e y_4 em η_2 foram colocadas como 1 para garantir a identificação do modelo. Ao se escrever o modelo de mensuração, temos as matrizes de covariâncias:

$$\Theta_{\delta} = \begin{pmatrix} \theta_{\delta_{11}} & 0 & 0 \\ 0 & \theta_{\delta_{22}} & 0 \\ 0 & 0 & \theta_{\delta_{33}} \end{pmatrix} \text{ e } \Theta_{\varepsilon} = \begin{pmatrix} \theta_{\varepsilon_{11}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_{\varepsilon_{22}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{\varepsilon_{33}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{\varepsilon_{44}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{\varepsilon_{55}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{\varepsilon_{66}} \end{pmatrix}.$$

- O **modelo estrutural** é escrito como $\eta = B\eta + \Gamma\xi + \zeta$ que, no exemplo, fica

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ b_{21} & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} \\ \gamma_{21} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix},$$

em que b_{21} é o coeficiente de regressão de η_1 em η_2 ; γ_{11} e γ_{21} são coeficientes de regressão de η_1 em η_2 , por ξ_1 .

Temos ainda duas matrizes que se referem às covariâncias entre η_1 em η_2 :

$$\Phi = (\phi_{11}) \text{ e } \Psi = \begin{pmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{pmatrix},$$

em que ϕ_{11} representa a variância de ξ_1 , ψ_{11} representa a variância de ζ_1 e ψ_{22} a variância de ζ_2 .

ANEXO 2 – IDENTIFICABILIDADE EM MODELOS DE EQUAÇÕES ESTRUTURAIS

Uma função é dita **não-identificável** se para dois valores distintos de um parâmetro (ou vetor de parâmetros) a função retorna o mesmo valor.

Para exemplificar a situação em que o modelo é identificado ou não-identificado, ao se observar a situação da Figura A2.1(a), tem-se 3 correlações entre as variáveis observadas A, B e D, o que gera três equações e aqui, diz-se que o modelo está identificado. Mas, em (b), tem-se quatro parâmetros a serem estimados a, b, c e d e apenas três equações (uma para cada par de variáveis AB, AC e AD).

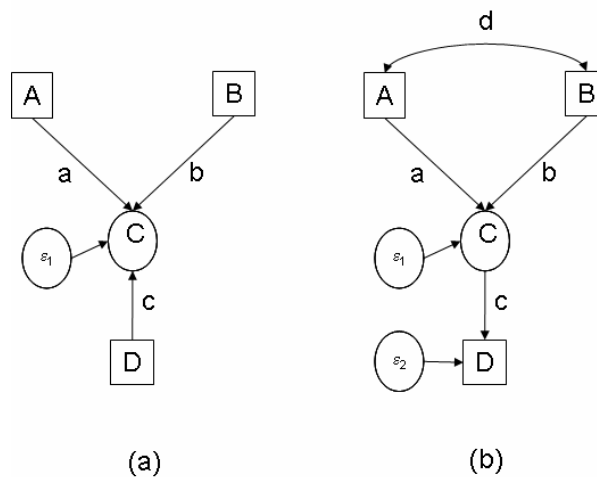


Figura A2.1. Exemplo de modelo identificado (a) e não-identificado (b).

Em MEE, a **identificabilidade** consiste em se verificar se a matriz de covariâncias Σ será gerada por apenas um único vetor de parâmetros. Se houver outro vetor de parâmetros que também gera a matriz Σ , então existe problema de identificabilidade. Mesmo que o modelo seja não-identificável, isto pode ser somente para alguns dos parâmetros do modelo e, assim, podem ser aplicadas restrições nos parâmetros para garantir a identificabilidade (Bollen 1989).

Uma outra forma de não-identificabilidade pode ser entendida quando temos mais de um conjunto de observações e que vão gerar o mesmo valor de verossimilhança (Hamilton 1994).

Existem algumas regras para se verificar a identificabilidade do modelo (Bollen 1989). Pela regra t , para haver identificabilidade do modelo é necessário que $t \leq \frac{1}{2}(p+q)(p+q+1)$, em que t é o número de parâmetros totais do modelo, p é o número de variáveis associados a x e q o associado a y . Essa regra é necessária, porém não suficiente. Como estas regras não são suficientes, alguns autores (Edward e Bogozzi 2000) acreditam que os problemas de não-identificação podem ser trabalhados no momento de se definir a parte estrutural do modelo.

ANEXO 3 – ESTIMAÇÃO PARA DADOS BINÁRIOS

Estimação - ADF

O método de distribuição assintótica livre (*ADF*) foi introduzido por Browne (1984) e foi importante por não assumir uma distribuição específica para as variáveis que tem sua covariância modelada. Para explicitar o efeito de variáveis categóricas, observa-se a função de ajuste a seguir:

$$F = [\hat{\Sigma} - \Sigma(\theta)]^T W^{-1} [\hat{\Sigma} - \Sigma(\theta)]$$

em que W é a função de pesos que é a matriz assintótica de covariâncias, ou seja, a matriz de covariâncias das variâncias e covariâncias estimadas a partir das variáveis observadas na amostra.

Para dados contínuos, a matriz W é calculada e o cálculo da sua inversa (W^{-1}) pode ser computacionalmente mais intenso dependendo do tamanho da matriz, ou seja, do número de variáveis observadas utilizado.

Para se construir a matriz W no caso de o estudo apresentar indicadores dicotômicos (ou ordinais), é necessário o uso de um modelo de pontos de corte ou limiares (*threshold model*), partindo-se da suposição de que as variáveis categóricas observadas proviriam, hipoteticamente, de variáveis contínuas com distribuição normal. Se houver razão para não considerar que a variável geradora tenha distribuição normal, os pontos de corte definidos pelo usuário podem até mesmo ser especificados. Para se estimar corretamente a matriz de covariâncias de indicadores dicotômicos ou ordinais, assume-se que essas variáveis contínuas tenham distribuição normal e, para tanto, estima-se a correlação entre as variáveis da seguinte forma: se as variáveis são categóricas, a correlação entre as variáveis latentes correspondentes é chamada de policórica; a correlação para duas variáveis dicotômicas é tetracórica e se uma variável for ordinal e a outra continua, ela é poliserial (Bollen 1989). Assim, como em Jöreskog & Sörbom (2002), recomenda-se que, se os dados são ordinais ou categóricos, o método de estimação a ser empregado

é o de mínimos quadrados ponderados (WLS) e a matriz de correlação utilizada deve ser a tetracórica ou policórica.

A idéia por trás da matriz de correlações tetracórica é a de que a variável observada y pode ser vista como uma observação parcial da variável latente contínua y^* . Então, este poderá ser escrito como: $y^* = v + \varepsilon$ em que v é um preditor linear e ε um termo de erro. Para variáveis contínuas, a resposta latente é simplesmente $y^* = y$. Quando se tem variáveis categóricas ordinais, podem-se escrever y da seguinte forma, em que τ_1, \dots, τ_s indicam pontos de corte (limiares):

$$y = \begin{cases} a_1 & \text{se } \tau_0 < y^* < \tau_1 \\ a_2 & \text{se } \tau_1 < y^* < \tau_2 \\ \dots & \\ a_s & \text{se } \tau_{s-1} < y^* < \tau_s \end{cases}, \text{ em que } \tau_0 = -\infty \text{ e } \tau_s = +\infty.$$

Com essa idéia pode-se ainda trabalhar com variáveis com censura intervalar, ordinais ou truncadas. Para variáveis dicotômicas, tem-se simplesmente:

$$y = \begin{cases} 1 & \text{se } y^* \geq \tau \\ 0 & \text{caso contrário} \end{cases}.$$

Em resumo, calculam-se as correlações policóricas, poliserial ou tetracórica para os pares de variáveis com distribuição conjunta não-normal, assumindo-se que eles são gerados por uma variável latente contínua, que tenha sua distribuição conjunta (*large sample joint distribution*) bivariada normal. Essas correlações são inseridas na estimação.

Para variáveis categóricas, o estimador *ADF* exige um tamanho de amostra extremamente grande (Finney e Disteffano 2006), o que é muitas vezes, incompatível com os estudos feitos em epidemiologia. Esse método produz estimativas eficientes dos parâmetros, mas tem limitações práticas como: o emprego de no máximo 25 variáveis que podem ser incluídas no processo de modelagem (Bollen e Long 1993). Essa indicação leva a que o processo de escolha de variáveis se torne mais seletivo, o que implica também em se restringir ou especificar mais o modelo causal. De acordo com Muthén e Muthén (2002), são várias as regras para se definir o tamanho da

amostra (tais como 5 a 10 observações por parâmetro, 50 observações por variável e não menos do que 100 indivíduos) ou, como diz Hox e Bechger (1998), que sugerem de “pelo menos 500 a 1000 indivíduos”. O tamanho da amostra necessária depende do número de variáveis, ocorrência de valores faltantes (*missing*), confiabilidade das variáveis e a força da relação existente entre as variáveis. Em um modelo de simulação de análise fatorial confirmatória, em que se empregam 2 variáveis latentes contínuas e 5 indicadores contínuos cada, foram obtidos 150 indivíduos para um poder de 0,81 para rejeitar a hipótese de que a correlação do fator é zero. Para indicadores não-normais, 265 indivíduos para o poder de 0,80 (Muthén e Múthen 2002). Aqui, os estimadores *ADF* não são adequados para variáveis ordinais, pois o estimador assume homocedasticidade dos erros (Bollen 1989). Outro problema apontado por Bollen (1989) ocorre na hipótese da matriz de covariâncias populacional não ser a mesma da matriz de covariâncias definidas no modelo.

Estimação – Satorra-Bentler

O método de Satorra-Bentler (*Satorra-Bentler scaling procedure*) produz uma correção no qui-quadrado que incorpora a curtose das variáveis. Como citado por Finney e Distefano (2006), diversos trabalhos examinam o uso desse método na situação de variáveis contínuas não-normais, enquanto que são poucos os trabalhos com variáveis categóricas ordinais.

Estimação - WLSMV

A seguir, detalha-se o método de mínimos quadrados ponderados, proposto por Muthén (1984), que será utilizado na estimação na análise da mortalidade fetal. Muthén (1984) propõe os seguintes passos para se fazer a estimação, que serão tratados a seguir: estimar os pontos de corte ou limiares (*thresholds*) populacionais de onde vem as variáveis observadas; estimar as correlações (policórica, tetracórica) populacionais dados os pontos de corte; e a estimação dos parâmetros do modelo dadas as correlações e pontos de corte obtidos.

Muthén (1984) ainda sugere que uma alternativa é utilizar escores fornecidos por um modelo probito ou logito para a variável categórica observada em um primeiro momento, e então o MEE é empregado baseando-se nesses escores, em um segundo

momento (Muthén 1993). Ao se assumir normalidade para os erros, pode-se utilizar o modelo probito:

$$F(\tau) = \Pr[\varepsilon \leq \tau] = \Phi^{-1}(\tau),$$

em que $\Phi^{-1}(\tau)$ é função de distribuição acumulada da normal padrão e τ é o ponto de corte.

Outra possibilidade é considerar os erros com distribuição logística. Tem-se

$$F(\tau) = \Pr[\varepsilon \leq \tau] = \frac{\exp(\tau)}{1 + \exp(\tau)}.$$

Apesar da diferença teórica entre o logito e o probito, na prática é difícil justificar o uso de uma e não de outra (Borooah 2001), a não ser por problemas de convergência por conta das técnicas de integração empregada no cálculo do logito (Muthén 1984).

O método GLS indica que a matriz de pesos é baseada em um estimador da matriz de covariâncias assintótica amostral analisada. Muthén (1984) desenvolveu e implementou dois estimadores robustos denominados WLSM (*weighted least squares mean adjusted*) e WLSMV (*weighted least squares mean and variance adjusted*) que, diferentemente da abordagem por ADF, não necessitam uma amostra de tamanho grande. O WLSM e WLSMV não são estimadores de mínimos quadrados generalizados porque são utilizados apenas alguns elementos da matriz de pesos W . A matriz WLSM utiliza a mesma matriz de covariâncias assintótica em que a diagonal da matriz de pesos é W com as estimativas de WLS. Assim, a covariância assintótica para o WLSMV é estimada por:

$$\hat{V}(\hat{\theta}_{\text{WLSMV}}) = n^{-1}(\Delta^T W^{-1} \Delta)^{-1} \Delta^T W^{-1} V_2 W^{-1} \Delta (\Delta^T W^{-1} \Delta)^{-1}, \text{ em que } \Delta = \partial \Sigma(\theta) / \partial \theta.$$

Para esta abordagem, o estimador de mínimos quadrados generalizados (GLS) pode ser utilizado para o modelo probito, o que permite solucionar correções como, por exemplo, no caso de heterocedasticidade (em modelos de regressão), usando a matriz de pesos W e a função de ajuste é escrita da seguinte forma:

$$F(\theta_{\text{WLS}}) = [\hat{\Sigma} - \Sigma(\theta)]^T W^{-1} [\hat{\Sigma} - \Sigma(\theta)].$$

Na estimação por WLS, W é estimado por um estimador consistente da matriz de covariância assintótica de $\hat{\Sigma}$ (Muthén 1984). A covariância assintótica desse estimador é:

$$\hat{V}(\hat{\theta}_{\text{WLSMV}}) = n^{-1}(\Delta^T V_2^{-1} \Delta)^{-1}, \text{ em que } \Delta = \partial \Sigma(\theta) / \partial \theta.$$

Assintoticamente, o estimador WLSMV não é tão eficiente quanto o WLS, mas em estudos de simulação (Muthén 1984) verificou que a inclusão de elementos de fora da diagonal da matriz W , como é feito para o estimador WLS, parece introduzir mais ruído do que melhorar a eficiência, na situação em que o tamanho de amostra não é tão grande. Satorra (1998) chamou essa abordagem de robusta.

Isso sugere que se possa escrever a variável categórica como uma transformação para uma função linear. Assim, modelos que misturam variáveis nominais, contínuas e ordinais podem ser unificados em um modelo linear generalizado (GLM). A implementação dessa idéia pode ser encontrada nos softwares Mplus (Muthén e Muthén 2004) e na biblioteca *gllamm* (*generalized linear latent and mixed models*) do software Stata, descrito por Skrondal (2004). Nesse contexto, os dois softwares incluem dados hierárquicos (multinível), efeitos aleatórios, dados faltantes, variáveis latentes (como modelo de classes latentes e modelos de crescimento latentes) e modelos de sobrevivência com tempos discretos. No *gllamm*, os modelos são especificados utilizando-se uma linguagem matricial ao contrário do software Mplus que utiliza uma linguagem mais simples, menos técnica. Um outro aspecto que diferencia os dois softwares é que o tempo de processamento do Mplus é bem menor que no *gllamm*, pois o Mplus evita utilizar o algoritmo de integração numérica ao fazer a estimação por máxima verossimilhança ou por mínimos quadrados generalizados. O *gllamm* não gera índices de ajuste diretamente, disponibilizando a verossimilhança para o cálculo de alguns índices.

Estimação - jackknife, bootstrap

Alternativas para se lidar com o problema de desvio da normalidade multivariada que afeta os erros-padrão é usar a estimação por máxima verossimilhança robusta ou o uso de técnicas de reamostragem como *jackknife* ou *bootstrap* e, assim, obter os erros-padrão dos parâmetros do MEE (Bollen 1993).

ANEXO 4 – MEDIDAS DE AJUSTE

Basicamente, o teste de hipóteses em MEE rejeita a hipótese de um modelo teórico se ajustar à matriz de covariâncias da amostra obtida. A hipótese nula é $H_0 : \Sigma = \Sigma(\theta)$, em que Σ é a matriz de covariâncias populacional das variáveis observadas, θ é o vetor que contém os parâmetros do modelo, e $\Sigma(\theta)$ é a matriz de covariâncias escrita em função dos parâmetros do modelo (Bollen 1989).

A partir de dois modelos, um restrito (da hipótese nula) e o outro completo (da hipótese alternativa), cada modelo gera uma medida de qui-quadrado de ajuste e a diferença entre os dois modelos é calculada como $D^2 = \chi_0^2 - \chi_1^2$, com graus de liberdade como a diferença entre os graus de liberdade dos modelos restrito (0) e completo (1). A hipótese nula H_0 é rejeitada se D^2 exceder o valor de qui-quadrado crítico. O teste de qui-quadrado é utilizado para se testar $H_0 : \Sigma = \Sigma(\theta)$ e, assim, o teste é feito para o modelo como um todo (Bollen 1984). Esse teste é válido ao se utilizar os estimadores de máxima verossimilhança, mínimos quadrados ponderados e não-ponderados. Pode-se, ainda, utilizar o teste de razão de verossimilhanças.

As medidas apresentadas a seguir são baseadas no qui-quadrado.

Índice de Tucker-Lewis (TLI)

$$TLI = \left[\frac{\frac{\chi_0^2}{gl_0} - \frac{\chi_1^2}{gl_1}}{\frac{\chi_0^2}{gl_0} - 1} \right],$$

em que N é o tamanho da amostra, gl é o número de graus de liberdade do modelo.

O TLI varia de 0 (sem ajuste) até 1 (ajuste perfeito)

Índice de ajustamento comparativo (CFI)

$$CFI = \frac{(\chi_0^2 - gl_0) - (\chi_1^2 - gl_1)}{\chi_0^2 - gl_0},$$

em que N é o tamanho da amostra, gl é o número de graus de liberdade do modelo.

O CFI varia de 0 (sem ajuste) até 1 (ajuste perfeito)

Raiz do erro quadrático médio de aproximação (root mean square error of approximation – RMSEA)

$$RMSEA = \sqrt{\frac{\chi_1^2}{(gl_1 - 1)(N - 1)}},$$

em que N é o tamanho da amostra, gl é o número de graus de liberdade do modelo.

O valor 0 de RMSEA indica ajuste perfeito.

Raiz do resíduo médio ao quadrado ponderado (Weighted Root Mean Square Residual - WRMR)

$$WRMR = \sqrt{\frac{1}{e} \sum_r \frac{(s_r - \hat{\sigma}_r)^2}{v_r}},$$

em que s_r é um elemento do vetor com as variâncias amostrais, $\hat{\sigma}_r$ é seu valor estimado, v_r é uma estimativa da variância assintótica de s_r , e é o número de parâmetros.

O WRMR é a diferença entre as matrizes de covariância observada e predita. O valor 0 indica ajuste perfeito.

ANEXO 5 – TABULAÇÕES DAS VARIÁVEIS OBSERVADAS

Tabela A5.1. Estatística descritiva das variáveis contínuas do estudo.

variável	freqüência	%
<i>união</i>		
sim (0)	339	71,1
não (1)	138	28,9
<i>tipo de habitação</i>		
alvenaria (0)	379	79,5
demais (1)	98	20,6
<i>classificação socioeconômica</i>		
A (1)	9	1,9
B (2)	64	13,4
C (3)	168	35,2
D (4)	206	43,2
E (5)	30	6,3
<i>ocupação do chefe de família</i>		
qualificado + semi-qualificado (0)	140	29,4
manual (1)	269	56,4
sem emprego (2)	68	14,3
<i>plano de saúde</i>		
sim (0)	155	32,5
não (1)	322	67,5
<i>planejamento da gravidez</i>		
planejada (0)	164	34,4
não planejada (1)	313	65,6
<i>tentativa de aborto</i>		
não (0)	405	84,9
sim, não tentou (1)	57	12,0
sim, tentou (2)	15	3,1
<i>reação da mãe</i>		
feliz (0)	312	65,4
infeliz/indiferente (1)	165	34,6
<i>reação do pai</i>		
feliz (0)	347	72,8
infeliz/indiferente (1)	130	27,2
<i>reação da família</i>		
feliz (0)	310	65,0
infeliz/indiferente (1)	167	35,0
<i>intercorrências</i>		
não (0)	304	63,7
sim (1)	173	36,3
<i>adequação de pré-natal</i>		
adequado (0)	316	66,3
não adequado (1)	161	33,7
<i>retardo de crescimento intra-uterino (RCIU)</i>		
não (0)	374	78,4
sim (1)	103	21,6

Tabela A5.2. Estatística descritiva das variáveis contínuas do estudo.

variável	freqüência	média	desvio-padrão	mínimo	máximo
anos completos de estudo do chefe da família	477	6,45	3,53	0,00	13,00
anos completos de estudo da mãe	477	7,36	3,25	0,00	13,00
renda familiar per capita (em salários mínimos)	477	1,82	2,51	0,01	22,22
densidade pessoa-cômodo	477	1,39	0,98	0,33	9,00

ANEXO 6 – COMITÊ DE ÉTICA EM PESQUISA**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE SAÚDE PÚBLICA
COMITÊ DE ÉTICA EM PESQUISA - CEP**

Av. Dr. Arnaldo, 715 - CEP 01246-904 - São Paulo - Brasil

São Paulo, 15 de setembro de 2001

Pelo presente, informo que o Comitê de Ética em Pesquisa da Faculdade de Saúde Pública da Universidade de São Paulo, **analisou e aprovou** as justificativas expostas pelos autores do projeto “Mortalidade perinatal na região sul do município de São Paulo”, ao entender que de acordo com o item III.3, g – o projeto deve “contar com o consentimento livre e esclarecido do sujeito da pesquisa e/ou seu representante legal”.

Porém, em virtude do teor das questões apresentadas aos sujeitos de pesquisa – utilização de drogas ilícitas, presença de violência - poderem, caso haja alguma forma de quebra de anonimato dos respondentes, causar riscos de estigmatização, formas de preconceitos ou problemas relacionados à esfera jurídica, aprovar que o consentimento livre e esclarecido seja, excepcionalmente, obtido de forma oral, mediante leitura do termo de responsabilidade do pesquisador, trazendo os objetivos da pesquisa, seus procedimentos e a finalidade, assim como a garantia de anonimato, sem a necessidade da existência de documentos que possam identificar as pessoas que, voluntariamente, se disponham a participar do estudo.

Atenciosamente

PAULO ANTONIO DE CARVALHO FORTES
Professor Associado
Coordenador do Comitê de Ética em Pesquisa da FSP-USP