

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA**



DISSERTAÇÃO DE MESTRADO

**RISCO DE ANOMALIA CONGÊNITA POR MISOPROSTOL
ATRAVÉS DO MODELO DE COX COM TRUNCAMENTO À
ESQUERDA**

Marília Canabarro Zordan

Orientador: Prof. Dr. Álvaro Vigo

Co-orientador: Profa. Dra. Tatiane da Silva Dal Pizzol

Porto Alegre, 2009

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA



DISSERTAÇÃO DE MESTRADO

**RISCO DE ANOMALIA CONGÊNITA POR MISOPROSTOL
ATRAVÉS DO MODELO DE COX COM TRUNCAMENTO À
ESQUERDA**

Marília Canabarro Zordan

Orientador: Prof.Dr. Álvaro Vigo
Co-orientador: Profa. Dra. Tatiane Dal Pizzol

A apresentação desta dissertação é exigência do Programa de Pós-graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul, para obtenção do título de Mestre.

Porto Alegre, Brasil.
2009

Z88r Zordan, Marília Canabarro

Risco de anomalia congênita por misoprostol através do modelo de cox com truncamento à esquerda / Marília Canabarro Zordan ; orient. Álvaro Vigo ; co-orient. Tatiane da Silva Dal Pizzol. – 2009.

75 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Epidemiologia. Porto Alegre, BR-RS, 2009.

1. Anormalidades congênitas 2. Misoprostol 3. Análise de regressão 4. Sensibilidade e especificidade 5. Epidemiologia I. Vigo, Álvaro II. Dal Pizzol, Tatiane da Silva III. Título.

NLM: QZ 45

Catálogo Biblioteca FAMED/HCPA

BANCA EXAMINADORA

Prof. Dr. Sotero Mengue, Programa de Pós-Graduação em Epidemiologia,
Faculdade de Medicina, Universidade Federal do Rio Grande do Sul

Prof. Dra. Patrícia Klarmann Ziegelmann, Departamento de Estatística,
Universidade Federal do Rio Grande do Sul

Prof. Dra. Lavínia Schüler-Faccini, Departamento de Genética, Instituto de
Bióciências, Universidade Federal do Rio Grande do Sul

AGRADECIMENTOS

Agradeço ao professor Álvaro Vigo por todo o apoio, paciência e dedicação investidos e imprescindíveis para este trabalho.

Muito obrigado especial para a Dra. Tatiane Dal Pizzol que teve contribuição fundamental para desenvolver desse trabalho.

E a todos aqueles que ao longo de todos os anos agüentaram me ouvir falando sobre esse assunto e manifestaram sua empatia.

SUMÁRIO

Abreviaturas e Siglas.....	6
Resumo.....	7
Abstract.....	8
1. APRESENTAÇÃO.....	9
2. INTRODUÇÃO.....	10
3. REVISÃO DA LITERATURA.....	13
3.1 Medicamentos e a Gestação.....	13
3.2 Medicamentos para induzir a menstruação.....	15
3.3 Anomalia Congênita.....	19
3.4 Análise Estatística.....	21
3.4.1 Métodos para análise de dados de sobrevivência.....	28
3.4.2 Razão de Riscos.....	33
3.4.3 Comparação de funções de sobrevivência.....	35
3.4.4. Modelando dados de sobrevida.....	38
3.4.5 Análise de sobrevida com truncamento à esquerda.....	42
4. OBJETIVOS.....	48
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	49
6. ARTIGO.....	53
7. CONCLUSÕES E CONSIDERAÇÕES FINAIS.....	66
8. ANEXO	
Projeto de Pesquisa.....	69

ABREVIATURAS E SIGLAS

HR – Hazard Ratio (Razão de Riscos ou Razão de Azares)

IMC – Índice de Massa Corporal

OMS – Organização Mundial da Saúde

RC – Razão de Chances

RR – Risco Relativo

RESUMO

Introdução: O uso de medicamentos no período entre a concepção e a confirmação da gravidez é um aspecto importante da epidemiologia reprodutiva, particularmente quando é usado com a finalidade de forçar a menstruação. No Brasil, o misoprostol tornou-se rapidamente popular pelo seu poder abortivo, porém sua taxa de eficácia não está claramente definida. Na maioria dos estudos com gestantes a entrada no estudo ocorre em tempos (idades gestacionais) diferentes. Assim, o tempo entre a concepção e o evento de interesse usualmente é truncado à esquerda e não considerar formalmente este aspecto na análise pode produzir vieses. Neste estudo é investigada a associação entre anomalia congênita e o uso do misoprostol, comparando três métodos de análise. **Métodos:** Coorte de 4251 mulheres com 20 anos de idade ou mais, arroladas entre a 21^a e a 28^a semana de gestação e acompanhadas até o parto, procedentes de serviços de pré-natal do SUS localizadas em seis capitais brasileiras. Foram comparadas a razão de azares produzida pela regressão de Cox com e sem truncamento à esquerda, e a razão de chances estimada pela regressão logística. **Resultados:** O uso de misoprostol foi identificado como fator de risco para anomalia congênita pelo modelo logístico (RC=2,9; IC95%: 1,12–7,44) e pelos modelos de Cox sem truncamento à esquerda (HR=3,3; IC95%:1,28-8,43) e com truncamento à esquerda (HR=4,8; IC95%:1,69-13,79), porém os preditores dos modelos multivariados são diferentes. **Conclusão:** A escolha do método de análise influencia os modelos finais e a magnitude da associação estimada. A característica de temporalidade incluída no modelo de Cox com truncamento à esquerda agrega informações importantes, principalmente, por se tratar de um desfecho que altera suas probabilidades de acontecimento com o desenvolver da gestação.

ABSTRACT

Introduction: The use of drugs in the period between conception and confirmation of pregnancy is an important aspect of reproductive epidemiology, particularly when used with the purpose of forcing menstruation. In Brazil, misoprostol quickly became popular for its abortive power, but its rate of effectiveness is not clearly defined. Most studies with pregnant women reports that they entering in the study in time (gestational age) different. Thus, the time between conception and the event of interest is usually truncated to the left and not formally consider this aspect in the analysis can produce bias. In this study we investigated the association between congenital anomalies and the use of misoprostol, comparing three methods of analysis. **Methods:** Cohort study of 4251 women aged 20 or older, enrolled between 21 and 28 weeks of gestation and followed until delivery, coming from services prenatal SUS located in six Brazilian capitals. We compared the hazard ratio produced by Cox regression with and without truncation to the left, and odds ratios estimated by logistic regression. **Results:** The use of misoprostol was identified as a risk factor for congenital anomalies by the logistic model (OR=2.9, 95%:1.12-7.44) and the Cox models without truncation to the left (HR=3.3, 95%:1.28-8.43) and left truncation (HR=4.8, 95%:1.69-13.79), but the predictors of the multivariate models are different. **Conclusion:** The choice of method of analysis influences the final models and the magnitude of the association estimated. A feature of temporality included in the Cox model with left truncation adds important information, especially since it is an outcome that changes your odds of developing the event of pregnancy.

1. APRESENTAÇÃO

Este trabalho consiste na dissertação de mestrado intitulada “**Risco de Anomalia Congênita por Misoprostol através do Modelo de Cox com Truncamento à Esquerda**”, apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, em 18 de dezembro de 2009. O trabalho é apresentado em três partes, na ordem que segue:

1. Introdução, Revisão da Literatura e Objetivos;
2. Artigo;
3. Conclusões e Considerações Finais.

O Projeto de Pesquisa está apresentado em anexo.

2. INTRODUÇÃO

Em uma reprodução bem sucedida, a variedade de eventos que pode acontecer entre a concepção e o nascimento de uma criança saudável é biológica e epidemiologicamente complexa. Problemas que surgem durante o curso do processo reprodutivo definem os desfechos adversos em estudos epidemiológicos sobre gravidez. A necessidade de uma atenção cuidadosa ao tempo dos eventos na epidemiologia reprodutiva aplica-se ao delineamento dos estudos e à análise dos dados. É necessário o reconhecimento explícito do tempo de ocorrência do desfecho, da exposição, do momento em que a gravidez passou ser observada, ou quando a mãe tornou-se consciente da sua gravidez (e daí, potencialmente consciente de seu término) ou quando chega para atendimento no sistema de saúde. (1)

No que diz respeito à farmacoterapia, para a maioria dos medicamentos, existem poucos dados sobre os efeitos no feto no momento em que o produto é comercializado. Essas informações geralmente provêm de estudos em animais, que nem sempre chegam a ser suficientemente esclarecedores sobre efeitos colaterais e a dosagem. Embora se tenha percebido que todo medicamento que causa teratogenicidade em humanos também causa efeitos similares em animais, alguns ainda são exceção, como é o caso do misoprostol (2,3). Dessa forma, apenas a experiência com o uso do medicamento em mulheres grávidas permite um julgamento adequado sobre o seu potencial teratogênico. Considerando que as gestantes são excluídas da maioria dos ensaios clínicos, por razões éticas, a avaliação da teratogenicidade dos medicamentos nos estudos pré-comercialização é limitada. Estudos epidemiológicos, tais como estudos de coorte e caso-controle, têm sido uma alternativa importante também para avaliação do risco teratogênico (3).

Para a mulher, o período entre a concepção e o diagnóstico da gravidez, por sua vez, pode estender-se por mais de um mês, aumentando o risco de uso inadvertido de medicamentos com potencial teratogênico. Conforme mostrado por Victora et al, em média, somente 66% das mulheres brasileiras procuram atendimento no primeiro trimestre da gravidez, não necessariamente no primeiro mês (3,4). Ainda, outros estudos apontam que a utilização de medicamentos e outros produtos com ação farmacológica para induzir a menstruação, quando está atrasada,

têm sido associada ao diagnóstico de gravidez ou à regularização do ciclo menstrual (3,5).

O misoprostol foi introduzido no Brasil em 1986 pelo laboratório Searle, com o nome comercial de Cytotec[®] (3). Embora licenciado para o tratamento de úlcera gástrica e duodenal, logo ganhou popularidade como abortivo (6,7). A bula do remédio, inclusive, possui um aviso de que o medicamento é contra-indicado durante a gravidez devido as suas propriedades abortifacientes (8). O produto foi comercializado sem restrições nos balcões das farmácias e drogarias até julho de 1991, quando o Ministério da Saúde proibiu a venda sem apresentação e retenção da prescrição médica.

Mesmo após muitos estudos terem sugerido que existe risco associado ao uso desse agente e da grande diminuição na sua administração, sabe-se que este fármaco continua sendo utilizado em alta escala, provavelmente para induzir aborto, principalmente em países onde esta prática é proibida, como é o caso do Brasil (3,5,9). O misoprostol se tornou uma alternativa popular e de largo uso para outros métodos empregados tais como chás, ervas, sondas, objetos perfurantes, entre outros, sendo as razões apontadas para o seu uso a privacidade, segurança e recusa de aborto em clínicas privadas para as gestantes. Porém, para o feto, é um dos métodos mais perigosamente invasivos para indução do aborto, conforme mostram alguns estudos realizados no país (6,7,10,11).

As taxas de falha associadas ao misoprostol em ensaios clínicos controlados têm variado, em média, de 5% a 10% (9). Se sob condições controladas proporcionadas pelos ensaios clínicos, o uso de misoprostol apresenta taxas de falha abortiva indesejáveis, nas situações em que o aborto ilegal é comumente realizado, sem assistência médica adequada e com emprego de doses inadequadas, o misoprostol pode não interromper a gestação, gerando ansiedade quanto ao risco para o feto (3,8).

Em relação ao tipo de análise estatística utilizada sabe-se que muitos estudos sobre aborto utilizam regressão logística para estimar a associação entre o desfecho e os fatores de exposição. Porém, este método não leva em consideração que as mulheres entram no estudo em tempos diferentes, uma vez que, em geral, são arroladas a partir de clínicas de pré-natal, isto é, após a concepção. Ou seja, as

mulheres entram no estudo em diferentes idades gestacionais e, mais importante, uma parte da população fonte é desconhecida por causa das gravidezes perdidas antes que houvesse o arrolamento. Assim, as mulheres que levaram a gestação até o nascimento têm maior oportunidade de entrar no estudo. Este tipo de dado é dito truncado à esquerda, e a análise de sobrevivência é mais recomendada por ser capaz de representar o tempo em observação. Como a entrada no estudo é diferenciada pelo desfecho e pode ser também diferenciada pelo status de exposição, o método analítico da regressão logística não permite tratar o truncamento à esquerda, podendo resultar em vieses na análise (12).

A análise de sobrevivência é capaz de representar cada um dos tempos de observação, e dessa forma, o truncamento à esquerda, considerando uma série de conjuntos de risco que incluem somente aquelas mulheres que entraram no estudo a tempo de ser observado o desfecho naquele particular tempo. A suposição é que as mulheres que tiveram adicionada a observação para uma determinada idade gestacional são representativas de todas as gravidezes sobreviventes para aquele período na população fonte (incluindo aquelas que nunca estiveram sob observação para aquela idade gestacional e aquelas que já estão sob observação) (12).

Dentro deste contexto, o presente trabalho pretende comparar empiricamente os resultados dos modelos de regressão logística, regressão de Cox e regressão de Cox com truncamento à esquerda, para estimar a associação entre o desfecho anomalia congênita com a exposição ao misoprostol, ajustadas e não ajustadas pelos outros fatores. O objetivo básico é analisar o impacto da escolha do modelo nos resultados e seus possíveis vieses, juntamente com a estimação das medidas de associação adequadas para cada modelo.

3. REVISÃO DE LITERATURA

3.1 Medicamentos e a gestação

O uso de medicamentos durante a gravidez exige precauções quanto à escolha do fármaco e dosagem apropriada. É importante entender como as mudanças fisiológicas que ocorrem durante a gravidez na unidade materno-feto-placentária podem acarretar variações nos processos de absorção, distribuição e eliminação dos medicamentos (13). Entretanto, o principal foco de atenção em relação ao uso de medicamentos na gestação diz respeito à interferência na formação, crescimento e desenvolvimento do embrião e do feto, com destaque para os efeitos teratogênicos, ou seja, o desenvolvimento de malformações (estruturais ou funcionais) dos órgãos fetais (2).

A cada ano, novos fármacos são fabricados e aprovados. Milhares de pessoas participam dos estudos clínicos pré-comercialização, porém, a maioria são homens e mulheres não-grávidas. Sendo assim, existe uma escassez de informações sobre os efeitos no feto no momento em que o produto é comercializado e os únicos dados provêm de estudos em animais. Embora se tenha percebido que todo medicamento que causa teratogenicidade em humanos também tem causado efeitos similares em animais, alguns ainda são exceção, como é o caso do misoprostol. Ainda, em alguns casos os efeitos teratogênicos só aparecem em animais quando administrados em altas doses. Portanto, embora os estudos em animais possam identificar efeitos de teratogenicidade, pode ser difícil extrapolar estes efeitos para os humanos (2). Dessa forma, apenas a experiência com o uso do medicamento em mulheres grávidas permite um julgamento adequado sobre o seu potencial teratogênico (3).

Considerando que as gestantes são excluídas da maioria dos ensaios clínicos, por razões éticas, a avaliação da teratogenicidade dos medicamentos nos estudos pré-comercialização é limitada. Estudos epidemiológicos, tais como estudos de coorte e caso-controle, têm sido uma alternativa importante não apenas na avaliação do risco teratogênico, mas também no estabelecimento da efetividade do medicamento utilizado em condições clínicas usuais, em populações específicas (3).

O período entre a concepção e o diagnóstico da gravidez, por sua vez, pode se estender por mais de um mês, aumentando o risco de uso inadvertido de

medicamentos com potencial teratogênico. Esta situação pode ser agravada quando a gravidez não é planejada ou por condições socioeconômicas e culturais da gestante, que podem retardar em até vários meses sua primeira consulta pré-natal. Victora e Barros (4) estimaram que no Brasil, em 1996, 13,2% das mães não fizeram pré-natal, variando de 4,9% na Região Sul a 25,2% na Região Nordeste. Idealmente, o pré-natal deveria começar já no início da gravidez para permitir a detecção e tratamento de complicações logo no início. Porém, das que fizeram pré-natal, cerca de dois terços realizaram sua primeira consulta no primeiro trimestre, variando de 79,7% no Sul a 51,9% no Nordeste. Ainda segundo os autores, o atendimento precário é provavelmente o maior problema para a pouca utilização do serviço, principalmente entre mulheres com níveis de risco gestacional alto e que não tem acesso a médicos privados ou seguro saúde.

Enquanto alguns autores sugerem evitar o uso de fármacos no primeiro trimestre da gravidez (14), outros advertem que tão importante quanto ter precaução no uso de medicamentos é ter parcimônia na aplicação da recomendação geral de não usar, principalmente quando o não uso pode privar a gestante de um tratamento efetivamente necessário (3,15). Em determinadas situações, a gestante pode ter uma condição de saúde que requer tratamento contínuo, tais como hipertensão, diabetes, asma, depressão grave e epilepsia, ou pode desenvolver distúrbios agudos induzidos ou não pela gestação, como náuseas, vômitos, doenças infecciosas respiratórias ou genitais, entre outros. O uso de remédios nessas situações comumente não pode ser evitado, sendo importante escolher o tratamento mais efetivo para a mãe e o feto. Nos casos em que as mulheres planejam engravidar ou só depois de algum tempo descobrem a gravidez, os médicos devem se certificar de que elas entenderam claramente a natureza e a magnitude do risco associado ao uso do medicamento, deixando para a futura mãe a decisão de se o potencial benefício justifica o risco para o feto e, assim, de levar ou não a gestação adiante (2,3).

Os estudos sobre o uso de medicamentos na gestação realizados em vários países mostram variações expressivas principalmente quanto a algumas intervenções consideradas de rotina durante a gestação (16). Em geral, o uso de medicamentos durante a gestação pode ser agrupado por suas finalidades e classificadas em profiláticas, terapêuticas e não-terapêuticas (3).

Entre os medicamentos de uso profilático, ou seja, remédios que objetivam a prevenção de alguma desordem, os mais utilizados são os sais minerais e as vitaminas. Ainda, pode-se citar o uso de ácido fólico e do sulfato ferroso, este especialmente na segunda metade da gravidez, para a prevenção da anemia, desordem muito comum entre as gestantes (3,17).

O grupo terapêutico é constituído pelos medicamentos destinados ao tratamento de distúrbios agudos ou crônicos, relacionados ou não à gestação. Como exemplos, podem ser citados os antieméticos e medicamentos para o tratamento de complicações associadas à gestação como o diabetes gestacional. Além disso, algumas doenças pré-existentes como hipertensão arterial sistêmica, diabetes e epilepsia exigem, normalmente, manutenção do controle farmacológico em curso no período anterior à gestação.

Por fim, o uso não-terapêutico é definido como o emprego de medicamentos com indicações diversas daquelas estabelecidas pelo fabricante ou com finalidades abusivas, tais como o uso de medicamentos com efeito estimulante ou depressor do sistema nervoso central e de medicamentos para diagnóstico ou interrupção da gestação (3).

Nas próximas seções serão abordadas questões relacionadas com a segurança de medicamentos utilizados com fins não-terapêuticos para indução a menstruação, em especial, o misoprostol.

3.2 Medicamentos para induzir a menstruação

A utilização de medicamentos e outros produtos com ação farmacológica para induzir a menstruação quando está atrasada tem sido associada ao diagnóstico de gravidez e à regularização do ciclo menstrual. O uso intencional de medicamentos para estimular a menstruação, tais como progestágenos, associados ou não a estrógenos, se tornou popular na década de 70. Nos anos seguintes, o misoprostol, um derivado das prostaglandinas, surge como alternativa para indução da menstruação. Mesmo após muitos estudos terem apontado para a possibilidade do risco associado ao uso desse agente e que sua administração tenha sido bastante diminuída, existem evidências de que o misoprostol continue sendo utilizado,

provavelmente, para causar o aborto, principalmente em países onde esta prática é proibida, como é o caso do Brasil (3,5).

O aborto é uma prática ilegal no Brasil e mulheres grávidas buscam nas farmácias e drogarias indicação de medicamentos que induzem o aborto ou sangramento uterino, para que então elas possam ir ao hospital para realizar uma curetagem de emergência. Um estudo publicado em 1991 e que foi realizado em Fortaleza sobre a venda de abortivos em 102 farmácias revelou que na maioria dos estabelecimentos visitados (83%) foi recomendado o uso de misoprostol (3,18).

Como mencionado, o misoprostol é um dos métodos abortivos mais perigosamente invasivos para o feto, conforme mostram alguns estudos realizados no país (6,7,10,11). Um deles, realizado em sete hospitais públicos do Rio de Janeiro em 1993, entrevistou 1603 mulheres admitidas com complicações relacionadas ao aborto, das quais 50% foram classificadas como casos de aborto induzido. Ainda nesse estudo, o misoprostol, isolado ou em combinação com outro método, foi citado por 57% das mulheres como método mais utilizado na tentativa de aborto. A maioria das 458 usuárias do misoprostol (84%) deste mesmo estudo tomou conhecimento sobre o medicamento através de amigos, parentes ou colegas e apenas uma pequena proporção (10%) ouviu sobre ele diretamente do farmacêutico. Ainda, 28% das mulheres escolheram o misoprostol porque era o método mais acessível, 14% porque era mais seguro que outros métodos e 13% porque era o mais barato (6). Uma pesquisa semelhante realizada em Fortaleza, Ceará (1991), revelou percentuais de uso do misoprostol ainda maiores entre mulheres admitidas na maternidade com diagnóstico de aborto induzido; 75% das 593 mulheres atendidas declararam ter usado o remédio (7). Em outro estudo realizado entre 1992 e 1993 (10), o uso de misoprostol foi relatado por 66% de um total de 2074 mulheres com diagnóstico de aborto provocado. Em um estudo realizado em Florianópolis (1993-1994), Santa Catarina, verificou-se que entre 141 casos de indução do aborto atendidos na emergência de uma maternidade, 50,4% foram provocados pelo uso isolado de misoprostol (11).

No Estudo Brasileiro de Diabetes Gestacional (EBDG), na primeira fase da coorte, o grupo arrolado foi formado por gestantes maiores de 20 anos entre a 21^a e 28^a semana de gestação selecionadas em clínicas de pré-natal. Por meio de uma

questão uso-orientada, 874 (14%) mulheres informaram ter utilizado alguma substância para induzir o fluxo menstrual durante a gravidez. As substâncias usadas com maior frequência foram chás com ervas medicinais (41,1%), combinações de estrogênios e/ou progestágenos (29,3%) e o misoprostol (15,2%) (5).

O misoprostol é uma prostaglandina sintética, análoga à prostaglandina E₁, ativa por via oral e eficaz na prevenção e tratamento de úlcera péptica causada por inibidores da prostaglandina sintetase (19). Foi introduzido no Brasil em 1986 pelo laboratório Searle, com o nome comercial de Cytotec[®] (3). Embora licenciado para o tratamento de úlcera gástrica e duodenal, logo ganhou popularidade como abortifaciente, conforme mostram dois estudos brasileiros de Costa (6) e Coelho (7).

Uma revisão cuidadosa dos estudos sobre a eficácia do misoprostol para tratamento da úlcera gástrica mostra que a única justificativa segura para a sua indicação em lugar de outros medicamentos é a prevenção da erosão gástrica em pacientes de risco tratados com antiinflamatórios não esteróides, particularmente em idosos (20). Porém, este medicamento mostrou-se, na prática, ser mais utilizado para fins de ginecologia e obstetrícia, devido aos seus efeitos uterotônicos. Na bula do remédio consta um aviso de que o misoprostol seria contra-indicado durante a gravidez por causa de suas propriedades abortifacientes (8). O fármaco tem o potencial de aumentar o tônus uterino, favorecendo o desenvolvimento de contrações, além de causar o amolecimento do colo do útero, facilitando a dilatação necessária à realização de curetagem, bem como outros procedimentos obstétricos (20). Além disso, a efetividade em abortos eletivos, indução do parto, evacuação do útero pós-morte fetal, hemorragia pós-parto e maturação cervical tem sido muito investigada, particularmente em países onde o aborto é legalizado (8).

No Brasil, o produto foi comercializado sem restrições nos balcões das farmácias e drogarias até julho de 1991, quando o Ministério da Saúde proibiu a venda sem apresentação e retenção da prescrição médica. Apesar dos esforços para limitar o uso de misoprostol, estudos conduzidos em Fortaleza (21) e em Florianópolis (11) entre 1992 e 1994 mostraram que o uso declarado do produto em tentativas de aborto continuava elevado. As estimativas do uso do misoprostol verificadas nos estudos citados devem estar, possivelmente, subestimadas pelo viés

de informação introduzido pelas mulheres, dada a condição ilegal do aborto no Brasil, e os aspectos morais e religiosos envolvidos (3).

Atualmente, a Portaria nº344, publicada pela Secretaria de Vigilância Sanitária do Ministério da Saúde, em vigor desde 1998, estabelece que a compra e uso do misoprostol seja restrita aos estabelecimentos hospitalares cadastrados junto à Agência Nacional de Vigilância Sanitária (ANVISA). Entretanto, veículos de comunicação de massa têm denunciado a venda clandestina do produto em farmácias e drogarias, no comércio ambulante e até mesmo pela internet (3).

O risco à saúde materna não é o único e principal problema associado ao uso indiscriminado de substâncias abortivas. Ensaios clínicos controlados demonstram que a falha abortiva com o uso do misoprostol em abortos médicos pode ultrapassar 10%, exigindo emprego de método abortivo adicional para a consumação do procedimento (8). Uma revisão sistemática de ensaios clínicos controlados (3) revelou que existem alternativas terapêuticas mais efetivas para a indução do aborto, tais como combinações com mifepristona, metrotexato ou tamoxifeno (8). Assim, deve-se evitar o uso isolado do misoprostol para indução do aborto médico no primeiro trimestre (8). Outra revisão sistemática realizada em 2006 cita que ao associar o misoprostol a mifepristona as taxas de falha abortiva podem atingir até 8% (22). Porém de acordo com os resultados encontrados por Marions (2006) esta taxa pode ser ainda maior, pois a taxa de aborto total ficaria apenas em torno de 88%-89% (23), ou seja, apresentando uma baixa eficácia da associação medicamentosa. Se sob condições controladas proporcionadas pelos ensaios clínicos, o uso de misoprostol apresenta taxas de falha abortiva indesejáveis, nas situações em que o aborto ilegal é comumente realizado, sem assistência médica adequada e com emprego de doses inadequadas, o misoprostol pode não interromper a gestação, gerando ansiedade quanto ao risco para o feto (3).

O risco de anomalias congênitas e de outros desfechos adversos da gravidez associado à exposição ao misoprostol ainda não está adequadamente estabelecido. Nos últimos 15 anos, relatos de casos publicados na literatura sugerem um possível efeito teratogênico do misoprostol, particularmente para o desenvolvimento da Sequência de Moebius (3). A Sequência de Moebius caracteriza-se por paralisia congênita e não progressiva dos nervos cranianos VI e VII, quase

sempre bilateral, produzindo uma aparência facial pouco expressiva e estrabismo convergente. Frequentemente outros nervos cranianos apresentam-se comprometidos, uni ou bilateralmente, determinando ptose palpebral, estrabismo divergente, surdez, distúrbios da sensibilidade nos territórios inervados pelo trigêmeo, disfonia e atrofia da língua, que podem ser verificados em diversas combinações (24).

Em um estudo de revisão sistemática e meta-análise realizado por Dal Pizzol et al (3), ao relacionar a exposição ao misoprostol e o desenvolvimento ou não de anomalias congênitas, em particular para a Sequência de Moebius, foi estimado que ao expor o feto ao misoprostol a chance de desenvolver a anomalia é 25 vezes a chance daqueles não expostos (3).

Schüler et al. (25) não conseguiram confirmar a hipótese de risco teratogênico do misoprostol ao comparar as taxas de anomalias congênitas maiores e menores nos filhos de 86 gestantes expostas ao medicamento, comparando com os filhos de outras 86 não expostas.

3.3 Anomalia congênita

Define-se como anomalia ou malformação congênita a anomalia estrutural presente no nascimento. Uma definição mais ampla seria a expressão “defeito congênito”, incluindo toda anomalia funcional ou estrutural do desenvolvimento do feto decorrente de fator originado antes do nascimento, seja genético, ambiental, ou desconhecido, mesmo quando o defeito não é aparente no recém-nascido e só se manifesta mais tarde (26).

Nas populações humanas em geral, muitos são os casos de anomalias congênitas, cada qual apresentando características próprias quanto à incidência, influência genética, gravidade e morbidade (27). A incidência geral dos defeitos congênitos na América Latina não difere, significativamente, daquela encontrada em outras regiões do mundo (26). Apesar dos fatores genéticos estarem implicados em praticamente todas as doenças, como resultado de sua interação com o ambiente, o papel relativo do componente genético poderá ser maior ou menor. Dentre as doenças nas quais o componente genético é preponderante, podem ser citadas as

monogênicas, individualmente raras, as cromossômicas e as multifatoriais, que são responsáveis por grande parte das anomalias congênitas e também por muitos problemas da vida adulta (26).

Muitas anomalias de origem genética e algumas com fatores de risco ambientais têm sido identificadas, mas a maioria das anomalias congênitas permanece com etiologia desconhecida. Muitas provavelmente têm causas genéticas e ambientais (28).

No que diz respeito à influência do hábito de fumar durante a gravidez, ainda não está consolidada na literatura sua relação com o aumento na prevalência das anomalias congênitas. No estudo The Danish National Birth Cohort (1997–2003), a exposição de gestantes ao tabaco durante as primeiras 12 semanas de gestação não evidenciou aumento na prevalência de anomalias congênitas (5,28%), comparada a prevalência em não expostas ao tabaco (5,11%) ($p=0,359$) (29).

A obesidade materna também tem sido considerada fator de risco para anomalia congênita (30,31). Alguns estudos estimam o peso materno antes da gravidez através do IMC. De uma forma geral, esse critério é bastante aceito, pois, por décadas, o IMC tem sido usado como um índice para obesidade capaz de substituir as medidas de tamanho e composição do corpo para prever o risco de uma ampla variedade de desfechos (32). Para alguns tipos específicos de anomalia o peso parece ser independente de outros fatores como idade materna, fumo, álcool, diabetes, não uso de vitaminas e história prévia de determinado defeito congênito (30).

O diagnóstico de diabetes melito (tipo 1 ou 2), antes mesmo da concepção, também aparece na literatura como fator de risco para anomalias congênitas (31,33,34,35). Segundo Ray et al., mulheres que recebem cuidados antes da concepção, e por isso tem o diabetes mais estabilizado, tendem a ter um menor risco dos filhos apresentarem problemas de formação (34). É importante ressaltar que muitos estudos realizados que indicam a relação entre obesidade materna e anomalia congênita são para tipos de anomalias diferentes e específicas, tais como defeitos no tubo neural, defeitos congênitos no coração e fissura orofacial (31,36). Outros estudos ainda apontam que o diabetes diagnosticado durante a gravidez (diabetes gestacional) pode também estar associado com alguns tipos de anomalias,

principalmente no sistema nervoso central, embora as evidências nesse caso sejam menos conclusivas (39).

Além disso, Anderson et al. relatam em seus resultados que, em consonância com outros estudos, a obesidade materna e o diabetes de todos os tipos (incluindo diabetes gestacional) não se associaram com aumento de risco para defeitos congênitos quando examinados sozinhos, como único fator de exposição. Porém, mães que apresentaram diabetes e obesidade conjuntamente apresentaram aumento no risco de desenvolver determinadas anomalias quando comparadas com mulheres não obesas e não diabéticas. (31).

Também tem sido descrita na literatura a associação entre o baixo peso ao nascer e a presença de anomalias congênitas mais graves (33,37). Uma das razões para que o baixo peso ao nascer continue sendo reportado e estudado na epidemiologia e na saúde pública é que este pode ser medido com excelente precisão e validade (38).

Independente de sua classificação como fator de risco demográfico, estatístico, social, pessoal ou identificação inicial, a idade materna tem sido considerada condição de risco, sempre que a gravidez ocorre fora da faixa etária ideal para a parturição, ou seja, nos extremos da vida reprodutiva, sejam essas gestações precoces ou tardias. (35,39).

Outros estudos que se referem apenas a anomalias congênitas específicas citam outros fatores de risco como histórico familiar (27,35), número de gestações, aborto prévio (27) e nascimento pré-termo (35,37).

3.4 Análise Estatística

Muitos estudos sobre aborto utilizam a regressão logística para analisar a relação entre o desfecho e os fatores de exposição. Porém, este método não leva em consideração que as mulheres entram no estudo em tempos diferentes, uma vez que, em geral, são arroladas a partir de clínicas de pré-natal, após a concepção. Sendo assim, as mulheres entram no estudo em diferentes idades gestacionais e, mais importante, uma parte da população fonte é desconhecida devido às gravidezes perdidas antes que houvesse o arrolamento. Esta situação atribui às mulheres que

levaram a gestação até o nascimento uma maior oportunidade de entrar no estudo. Este tipo de dado é chamado de “truncado à esquerda” e, nesses casos, a análise de sobrevivência é mais recomendada por ser capaz de representar o tempo em observação. Além disso, permite incorporar informações para os grupos de riscos definidos nesses tempos, pois o método analítico comum da regressão logística não permite tratar o truncamento à esquerda, podendo resultar em vieses na análise (12).

A variedade de eventos que podem acontecer entre a concepção e o nascimento de uma criança saudável é biológica e epidemiologicamente complexa. Problemas que surgem durante o curso do processo reprodutivo definem os desfechos adversos em estudos epidemiológicos sobre gravidez (1).

Com o esboço dos eventos normais, pode-se considerar desvios que resultam em resultados adversos. Algumas questões metodológicas gerais no estudo da saúde reprodutiva serão primeiramente abordadas. Devido à decisão de reproduzir ser pessoal, alguns desfechos, por exemplo, concepção ou falha ao conceber, são observados somente em alguns casais em “risco” de concepção. Alguns dos eventos seguintes só podem ser observados depois que alguns pré-requisitos tenham sido encontrados: uma concepção de sucesso é necessária para se ter um feto em risco de ser abortado espontaneamente ou nascer pré-termo. A fim de caracterizar acuradamente a frequência dos eventos, existe a necessidade de uma especificação clara das unidades (pessoas, fetos ou períodos de tempo) que realmente estão em risco, o que pode ser desafiante (1).

A natureza temporal do processo reprodutivo é um tema recorrente. O tempo de acontecimento dos eventos é frequentemente ambíguo, justamente porque alguns desfechos não podem ser observados no tempo em que eles ocorreram. Por exemplo, anomalias congênitas frequentemente têm sua etiologia nas primeiras semanas de gestação, mas o feto deve sobreviver até algum ponto da gestação para uma anomalia ser diagnosticada. Ou ainda, a formação da placenta no início da gravidez pode ser um fator chave na influência do risco de nascimento pré-termo. Alguns desfechos são definidos como falhas de alguma coisa que ocorreu no tempo certo. Por exemplo, quando o tubo neural não fecha completamente, resulta em espinha *bífida*. Alguns resultados são definidos por eventos normais que ocorrem no tempo errado, como, por exemplo, o parto pré-termo, formalmente definido somente

pela idade gestacional em que o parto de um feto vivo ocorre. Finalmente, a manifestação de um evento pode não ser identificada instantaneamente: várias semanas podem passar entre a morte de um feto e o reconhecimento de um aborto ou, ainda, várias semanas e até meses podem transcorrer até que uma anomalia possa ser detectada (1).

A necessidade de uma atenção cuidadosa ao tempo dos eventos na epidemiologia reprodutiva aplica-se ao delineamento dos estudos e à análise dos dados. Atenção especial é recomendada para a definição do tempo do desfecho (por exemplo, idade gestacional na hora do parto), da exposição (sempre que possa variar ao longo do tempo), e do tempo em que a gravidez passou a estar sob observação: quanto ao momento em que a mãe tomou conhecimento da sua gravidez (e então, potencialmente consciente de seu término) e, também, com respeito ao momento no qual a gravidez passa a ser acompanhada pelo sistema de saúde (e, portanto, pelo investigador que estuda mulheres arroladas em serviços de pré-natal). Em muitos casos, a análise apropriada requer o uso de modelos estatísticos que permitam considerar o número de pessoas-tempo em risco, como, por exemplo, modelos de riscos proporcionais de Cox. Hertz-Picciotto e colaboradores (2002) sugerem a aplicação do modelo de Cox para estudos de aborto espontâneo considerando a entrada para cuidados de pré-natal em idades gestacionais diferentes. Tais métodos também se aplicam ao estudo de eventos que ocorrem durante o curso da gestação, incluindo, perda, nascimento pré-termo, ou início de complicações (1).

Análise de sobrevivência é o nome para ao conjunto de técnicas estatísticas usadas para descrever e quantificar dados que informam o tempo até a ocorrência de um evento (40). Este termo é usado para descrever a análise dos dados em que desfecho é o tempo decorrido desde um tempo de origem até a ocorrência de algum particular evento, abandono ou final do estudo. O tempo até realização de algum evento é genericamente chamado de “tempo de falha”, podendo representar a falha de um componente eletrônico, uma máquina, morte ou infarto de algum paciente, ou o tempo de trabalho de um indivíduo após um período de desemprego (41).

Em pesquisa médica, o tempo de origem freqüentemente corresponde ao momento do recrutamento para o estudo, do diagnóstico de uma particular condição,

a data de aleatorização, o começo de um tratamento ou a ocorrência de alguma condição adversa. Se o evento de interesse no estudo é o óbito do participante ou paciente, o dado resultante será literalmente o tempo de sobrevivência. Porém, dados podem ser obtidos de uma forma similar quando o ponto final não é fatal, como a recorrência de um sintoma (42).

Existem características particulares de dados sobre “o tempo até o acontecimento de um evento de interesse” que não permitem a utilização de métodos estatísticos padrões, como a regressão logística. Primeiro porque, em geral, os dados não são simetricamente distribuídos. Ao se construir um histograma a partir dos tempos de sobrevivência de um grupo de indivíduos similares este tenderá ser positivamente assimétrico, isto é, o histograma terá uma longa “cauda” à direita do intervalo que contém o maior número de observações. Como consequência disso, não é correto assumir que os dados desse tipo têm distribuição normal. Esta dificuldade poderia ser resolvida primeiramente transformando os dados para tentar dar uma simetria maior à distribuição. Porém, uma aproximação mais satisfatória é adotar um modelo de distribuição alternativo para os dados originais (42).

Outra característica importante desse tipo de dados é que os indivíduos podem estar sujeitos a uma forma incompleta de observação (censura), podendo tornar os métodos padrões (regressão logística, por exemplo) inapropriados (41,42). Para alguns indivíduos o tempo completo T até a falha pode ser observado, enquanto que para outros só se sabe que T excede algum tempo de censura c (41). Estudos de coorte são usualmente delineados com um específico ponto final em mente, em geral, a ocorrência ou não de um determinado desfecho (doença). Durante o decorrer do seguimento um sujeito desenvolve a doença, ou não. Se a doença ocorre, o seguimento no estudo pode cessar para aquele indivíduo ou pode continuar para observar outros desfechos. Para aquele indivíduo o tamanho do seguimento é definido como o tempo desde o início do seguimento até o diagnóstico da doença, independente do que acontecer depois. Se a doença não se desenvolveu, o seguimento continua até o sujeito não ser mais observado ou até o término do estudo, o que acontecer primeiro. Neste caso, o tempo de seguimento é definido como o tempo do início do seguimento até a precedência de um dos dois eventos (43).

Outra forma de informação incompleta sobre os tempos de acompanhamento ocorre quando há truncamento, sendo importante diferenciar da censura.

O tempo de sobrevivência de um indivíduo é dito ser censurado quando o momento em que ocorre o evento não foi observado para aquele indivíduo (42). Existem três categorias principais de censura: à direita, à esquerda e intervalos de censura (44).

A censura à direita parte do pressuposto que o indivíduo já faz parte da amostra, estando então dentro do período de observação, que começa em τ_0 . Por exemplo, em estudos longitudinais como uma coorte, o tempo exato de sobrevivência só é conhecido para aqueles indivíduos que mostraram o evento de interesse. Para todos os outros (isto é, aqueles que não manifestaram o evento até o final do período de observação ou aqueles que foram perdidos) pode-se dizer que eles não manifestaram o evento de interesse durante o período do estudo ou em que estiveram em observação (40). Então, a censura à direita pode ocorrer, basicamente, de três maneiras. Primeiro, durante o período de observação ocorre uma perda de seguimento do indivíduo, ou seja, no momento de avaliação (visita clínica, por exemplo) o sujeito não comparece (por motivo de mudança ou impossibilidade de contato, por exemplo). Dessa forma, a única informação disponível sobre o referido participante é que até o momento da última avaliação o indivíduo ainda estava vivo ou saudável. Outra possibilidade é que o tempo de observação do estudo acabe e o sujeito não tenha manifestado o evento de interesse até então. Por fim, tem-se o caso em que o indivíduo manifesta uma característica que se sabe não ter relação com o fator em exposição em estudo, mas que por isso, precisa ser excluído. Por exemplo, se o desfecho de interesse é morte por ataque cardíaco e a pessoa simplesmente sofre um acidente de carro. Nessas situações, para um paciente que entrou no estudo no tempo τ_0 e não se conhece o tempo exato em que ele manifesta o desfecho, sabe-se apenas que ele continua vivo até o tempo $\tau_0 + c$, onde c é chamado o tempo de sobrevivência censurado. O tempo de censura pode variar de indivíduo para indivíduo e a única coisa que se sabe é que ele é menor que o real tempo da eventual ocorrência do evento que é desconhecido (42,44).

A censura à esquerda ocorre quando o verdadeiro tempo até a ocorrência de um evento é menor do que o tempo observado. Ou seja, o indivíduo é arrolado para um estudo e seguido por um período até a primeira avaliação e, neste momento, já apresenta o evento de interesse. Conseqüentemente, o real momento em que o evento aconteceu é desconhecido, e menor do que o tempo observado. A censura à esquerda é menos comum do que a censura à direita.

Por fim, uma censura intervalar ocorre quando se sabe que o indivíduo manifestou o evento de interesse dentro de um intervalo de tempo específico. Basicamente, pode-se dizer que este tipo de censura ocorre quando o indivíduo que está sendo observado manifesta o desfecho no intervalo de tempo entre duas avaliações, de tal forma que o tempo de início e o tempo de ocorrência do evento são desconhecidos (42).

Uma segunda característica que pode estar presente em alguns estudos de sobrevivência, algumas vezes confundido com censura, é o **truncamento**. O truncamento surge quando a amostragem de uma população é incompleta, ou seja, esse tipo de estudo não inclui aqueles sujeitos que não “sobreviveram” (ficaram sem manifestar o evento) tempo suficiente para que o evento fosse observado. Pode-se dizer então, que o truncamento é uma característica da amostra, em vez de individual. Por exemplo, ao selecionar os indivíduos, o investigador poderia nem estar ciente da existência daqueles que já manifestaram o desfecho num período anterior. Para dados truncados, somente indivíduos que experimentaram algum evento (não necessariamente o de interesse) são observados pelo investigador. O evento pode ser alguma condição que deve ocorrer **a priori** do evento de interesse, tal como a exposição a um fator de risco, ou seja, ocorrência de um fator intermediário antes do desfecho. O tipo mais comum de truncamento, à esquerda, ocorre quando sujeitos entram no estudo em períodos aleatórios (não necessariamente o tempo de origem, τ_0) e são seguidos a partir desse tempo de entrada “atrasado” até que o evento ocorra ou o sujeito seja censurado à direita. Neste caso, todos os sujeitos que experimentaram o evento de interesse antes deste tempo de entrada atrasado não serão conhecidos (nem selecionados), pelo experimentador. Note que, ao contrário da censura à esquerda, onde se tem uma informação parcial dos indivíduos que experimentaram o evento de interesse antes de

um determinado período ao entrar no estudo, estes indivíduos nunca serão incluídos no estudo (44,45).

A Figura 1 ilustra vários históricos de eventos completos e incompletos. O eixo horizontal indica um calendário de tempo iniciando em τ_0 . O período de observação é usualmente um período finito, com o início e o fim marcados por τ_0 e τ_1 , respectivamente. Por conveniência, o tempo de início da janela de observação é assumido ser *zero*. A letra minúscula “*e*” no final da linha indica o tempo exato em que a ocorrência de um evento é observada. A letra minúscula “*s*” que aparece a esquerda do final da linha indica que o tempo inicial é conhecido. A letra minúscula “*c*” indica que o sujeito é censurado naquele ponto. A letra minúscula “*t*” indica que o sujeito é truncado.

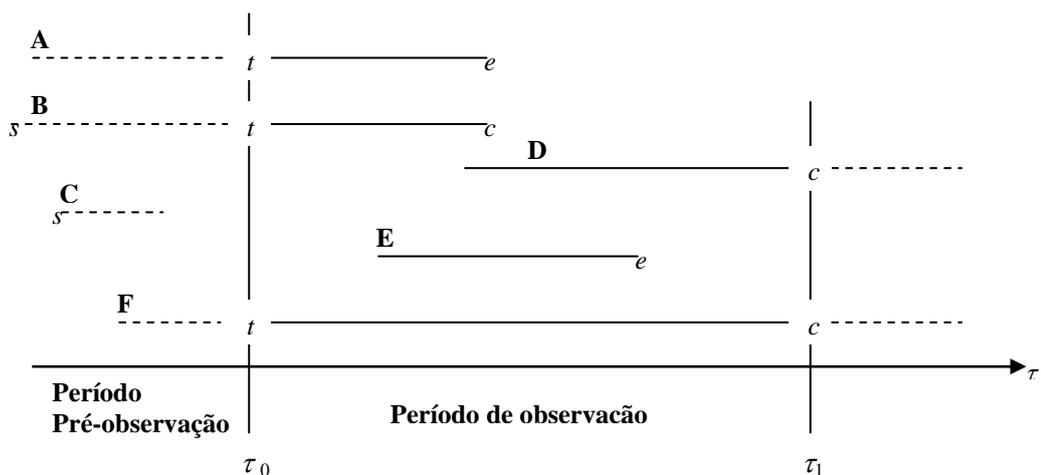


Figura 1: Históricos de eventos completos e incompletos.

Sendo assim, a partir da figura acima se pode dizer que os **sujeitos B, D e F** são censurados à direita; **B** é censurado pela perda de seguimento e **D e F** pelo término do período de observação. Os tempos dos **sujeitos A, B, C e F** podem ser ditos truncados à esquerda, pois foram expostos ao fator de risco por algum tempo antes de entrarem em observação. O **sujeito E** é o único que tem o histórico de eventos completo descrito na Figura 1. Dentre os sujeitos observados, os únicos que manifestaram o desfecho ou evento de interesse foram os sujeitos **A e E** (45).

Contextualizando o gráfico acima como se fossem 6 gestantes acompanhadas ao longo do tempo, poderiam ser realizadas as seguintes interpretações:

- 1) Não é possível saber o desfecho das gravidezes B, D e F; B por não ter dado prosseguimento no estudo e D e F pelo fato do estudo ter terminado antes do parto;
- 2) As gestantes A, B, e F teriam entrado no estudo após algum tempo de gestação, mas não imediatamente após a concepção. A gestante C nem teria entrado no estudo, pois a gravidez foi terminada antes do início do estudo;
- 3) A gestante E foi acompanhada desde o início da gestação até o momento do parto; e;
- 4) A gestante A, apesar de apresentar truncamento à esquerda foi acompanhada até a observação do desfecho.

3.4.1 Métodos para análise de dados de sobrevivência

Considere um estudo longitudinal, por exemplo, de coorte, conduzido durante um dado período de tempo $[\tau_0, \tau_1]$, onde τ_1 é a data de término do estudo. Seja τ' um tempo fixo tal que $\tau_0 < \tau' \leq \tau_1$ e suponha que os sujeitos são recrutados para o estudo num tempo base entre $[\tau_0, \tau']$, pois, tipicamente, os sujeitos não são todos recrutados ao mesmo tempo. Admita, também, que o seguimento comece imediatamente após o recrutamento. Este método de arrolamento é referido como entrada escalonada porque nem todos os membros da coorte são colocados sob observação ao mesmo tempo. Como resultado da entrada escalonada, os sujeitos inevitavelmente têm diferentes tempos de observação máxima, e ainda há o problema de que após o recrutamento alguns pacientes podem ser perdidos até começar o estudo (42,43). Por exemplo, alguém recrutado no tempo τ_0 terá um tempo de observação máximo de $\tau_1 - \tau_0$, enquanto que um indivíduo recrutado no tempo τ' terá um tempo máximo de observação de $\tau_1 - \tau'$. Mesmo se nenhum sujeito deixar de ser observado, a entrada escalonada e a variação no tempo máximo de observação resultarão em sujeitos com tamanhos diferentes de tempo de seguimento (43).

Para um dado indivíduo, considere t como sendo o tempo de sobrevivência e defina uma variável indicadora como: $\delta=1$ se o desfecho de interesse

é observado e $\delta=0$ em caso contrário. Quando $\delta=0$, se diz que t é um tempo de sobrevivência censurado, e quando $\delta=1$ que t é não censurado. Dessa maneira, o desfecho para cada sujeito é dicotômico – isto é, censurado ou não. Dados de sobrevivência para cada sujeito podem ser escritos de uma forma compacta como um vetor da forma (t, δ) , que se refere como uma observação. É comum dizer que uma observação é censurada ou não de acordo com o valor $\delta=0$ ou $\delta=1$, respectivamente (43).

Em geral, os estudos podem envolver dois tipos de tempo: o tempo do estudo em si e o tempo de sobrevivência, uma vez que os tempos de entrada são diferentes para cada indivíduo. Se for possível assumir que o recrutamento e desfechos são independentes do tempo de estudo, é apropriado igualar a dimensão de todos os tempos de entrada, isto é, todas as linhas de seguimento são colocadas no início e começam no mesmo ponto (43). Caso contrário deverá se levar em conta à existência de uma falta de informação à esquerda para alguns indivíduos.

Função de Sobrevivência e Função de riscos

Na teoria estatística da análise de sobrevivência, o tempo de sobrevida é considerado como uma variável aleatória contínua não negativa denotada por T (42,43). Assim, um tempo de sobrevivência t é uma realização de T e, como para qualquer variável aleatória contínua, T tem uma função densidade de probabilidade associada $f(t)$ (43). A função de distribuição de T é dada por $F(t) = P(T < t)$ e representa a probabilidade de que o tempo de sobrevida de um indivíduo seja menor que algum valor t (42). Na análise de sobrevivência é geralmente mais conveniente caracterizar T em termos de duas outras funções, chamadas de função de sobrevivência $S(t)$ e função de risco $h(t)$. A função de sobrevivência é definida por $S(t) = P(T \geq t) = 1 - F(t)$, isto é, $S(t)$ é a probabilidade de sobrevivência até ou além do tempo t (43). A função de sobrevida pode, portanto, ser usada para representar a probabilidade de que um indivíduo sobreviva do tempo de origem até algum tempo além de t (42). Suponha que o tempo amostrado de T é $[0, \tau]$, onde τ é o tempo máximo de sobrevivência possível de acordo com o delineamento do estudo. Por definição $S(0) = 1$, que significa que a coorte inteira está viva em t_0 . Como a intuição

sugere, $S(t)$ é uma função não-crescente de t , tal que para $t_1 < t_2$ implica em $S(t_1) \geq S(t_2)$ (43).

Para um tempo t arbitrário, mas fixo, considere um número positivo ε , pequeno. O intervalo $[t, t + \varepsilon)$ é o intervalo de tempos de sobrevivência maiores ou iguais a t e estritamente menores que $t + \varepsilon$. A probabilidade de morrer no intervalo $[t, t + \varepsilon)$, é $S(t) - S(t + \varepsilon)$, uma quantidade que se aproxima de zero à medida que ε se aproxima de zero (43).

Com t e ε como definidos anteriormente, a probabilidade condicional de morrer em $[t, t + \varepsilon)$, dado que sobreviveu até t , é

$$Q_\varepsilon(t) = \frac{S(t) - S(t + \varepsilon)}{S(t)}.$$

Uma vez que T é definida como uma variável aleatória contínua é comum considerar que a probabilidade de falha em qualquer tempo exato é zero. Assim, apenas quando se considera a probabilidade de morrer em um intervalo de tempo é que uma probabilidade diferente de zero pode ser obtida. Por esta razão, algumas vezes se refere a $f(t)$ e $h(t)$ como probabilidades “instantâneas” (43). Por exemplo, para $h(t)$, o risco instantâneo é assim chamado porque usa o número de falhas entre os tempos t e $t + \varepsilon$ dividido pelo número de pessoas em risco no tempo t , dividido por ε ; isto resulta na proporção de indivíduos presente no tempo t que morre ou falha por unidade de tempo (40).

Pode-se, então, caracterizar $S(t)$, $F(t)$ e $h(t)$ como segue: $S(t)$ é a probabilidade de que um indivíduo vivo no tempo 0 sobreviva além do tempo t ; $F(t)$ é a probabilidade instantânea por unidade de tempo que um indivíduo vivo no tempo 0 morra até o tempo t ; e $h(t)$ é a probabilidade instantânea por unidade de tempo que um indivíduo vivo no tempo t morrerá no “próximo instante” (43). Como exemplo das medidas “por unidade de tempo”, se o tempo de sobrevivência é medido em dias, então $h(t)$ é a probabilidade aproximada de que um indivíduo, que está vivo no dia t , morra no dia seguinte. Por esta razão, a função de risco é frequentemente interpretada apenas como o risco de falha no tempo t (42). Uma identificação útil da relação das três funções é $h(t) = \frac{f(t)}{S(t)}$ (43).

Uma vez coletados os tempos dos eventos (considerando as censuras), o primeiro passo é tentar descrevê-los, usualmente na forma gráfica como uma curva de sobrevivência. Visualizações ajudam a ter uma noção dos padrões temporais, além de também ajudar a identificar a forma de distribuição dos dados. Se os dados são consistentes com uma distribuição paramétrica, então os parâmetros podem ser estimados para descrever eficientemente o padrão de sobrevivência e inferências estatísticas podem ser baseadas na distribuição escolhida. Métodos não-paramétricos são usados quando distribuições teóricas não se ajustam aos dados. Um dos métodos não paramétricos que descrevem dados relacionados com o tempo de ocorrência é o método de Kaplan-Meier (40).

Para estimar a função de sobrevida, suponha primeiramente uma amostra simples dos tempos de sobrevida, onde nenhuma das observações seja censurada. A função de sobrevida $S(t)$, como definida anteriormente, é a probabilidade que um indivíduo sobreviva por um tempo maior ou igual que t . Esta função pode ser estimada por uma função de sobrevida empírica, dada por

$$\hat{S}(t) = \frac{\text{Número de indivíduos com tempo de sobrevida} \geq t}{\text{Número de indivíduos no grupo de dados}}.$$

Equivalentemente, $\hat{S}(t) = 1 - \hat{F}(t)$, onde $\hat{F}(t)$ é a razão do número total de indivíduos vivos no tempo t pelo número de indivíduos no estudo. Note que a função de sobrevida empírica é igual a 1 para valores de t antes do primeiro tempo de falha, e zero depois do tempo final para falhas (42).

A função de sobrevida estimada $\hat{S}(t)$ é assumida constante entre dois tempos de falha adjacentes, e então um gráfico de $\hat{S}(t)$ como função de t é uma “função escada”. A função decresce imediatamente após cada tempo de sobrevida observado (42).

Porém, este estimador empírico da função de sobrevida não pode ser usado quando existem observações censuradas. A razão para isso é que o método não permite utilizar informação proveniente de um indivíduo que o tempo de sobrevida é censurado antes do tempo t para ser usado na estimação da função de sobrevida em t . O método não-paramétrico de Kaplan-Meier é usualmente utilizado para estimar a função de sobrevida na presença de censura (42).

Curva de sobrevivência de Kaplan-Meier

O método de Kaplan-Meier é frequentemente usado para estimar a função de sobrevivência quando existem dados censurados (46). O método é baseado nos tempos individuais de sobrevivência e assume que a censura é independente do tempo de sobrevivência, isto é, a razão para uma observação ser censurada não está relacionada com a causa da falha (40). Além disso, ao construir uma série de intervalos para avaliar a sobrevida, o método só permite que para aqueles intervalos designados apenas um tempo de falha esteja contido em cada intervalo, e se considera que esta ocorreu no início do intervalo (42).

Considere uma coorte envolvendo r indivíduos cujas observações são denotadas por $(t_1, \delta_1), \dots, (t_i, \delta_i), \dots, (t_r, \delta_r)$. Assume-se que $t_i > 0$ para todo i , o que assegura que cada membro da coorte é seguido por pelo menos um pequeno espaço de tempo. Neste momento, os tempos de sobrevivência não censurados, isto é, aqueles em que o evento de interesse ocorreu, serão referidos como tempo de falha. Suponha que entre os r tempos de sobrevivência existam J tempos distintos ordenados de falha $\tau_1 < \dots < \tau_j < \dots < \tau_J$, $\tau_0 = 0$ e o máximo tempo de sobrevivência observado τ_{J+1} , isto é, $\tau_{J+1} = \max(t_1, t_2, \dots, t_r)$. O método de Kaplan-Meier particiona o período de seguimento em $J+1$ intervalos usando os tempos de falha como pontos de corte: $[\tau_0, \tau_1), [\tau_1, \tau_2), \dots, [\tau_j, \tau_{j+1}), \dots, [\tau_{j-1}, \tau_j), [\tau_J, \tau_{J+1}]$. Note que nenhum intervalo começa em um tempo de censura e o último intervalo contém o tempo τ_{J+1} (43). Em algumas aplicações serão consideradas censuras no tempo τ_{J+1} , devido àqueles sujeitos que sobreviveram (não experimentaram o evento de interesse) ao final do estudo. Se τ_{J+1} é um tempo de falha, então $\tau_J = \tau_{J+1}$ e o último intervalo reduz-se a um único ponto τ_J . Seja a_j o número de falhas em τ_j e c_j o número de observações censuradas no j -ésimo intervalo ($j = 0, 1, \dots, J$). Por definição, $a_0 = 0$ (43).

Como mencionado, o gráfico de $\hat{S}(t)$ é uma função escada onde as probabilidades de sobrevida estimadas são constantes entre tempos de falha adjacentes e decrescem a cada tempo de falha (40,42,43). O gráfico que considera o estimador de Kaplan-Meier de $S(t)$ no eixo vertical como uma função do tempo (no

eixo horizontal), é chamado de curva de sobrevivência de Kaplan-Meier. O gráfico começa em 1 (porque a probabilidade de sobreviver além do tempo 0 é 1) e termina em torno de zero. Se existirem sujeitos que sobreviveram além do tempo de estudo, então a curva de sobrevivência não se dirige para o zero, mas permanece na horizontal. Como a curva não é alterada nos intervalos nos quais não ocorrem eventos, a curva só pode ser estimada para os tempos que ocorrem eventos. O tempo médio para um evento é estimado pela área sob a curva da função de sobrevivência. Se o maior tempo de sobrevida observado é um tempo de falha, então a função de sobrevivência tende à zero naquele tempo e sua estimativa é finita. Do contrário, a média não pode ser estimada. Esta é a razão pela qual o tempo mediano de sobrevivência é usado mais freqüentemente que o tempo médio de sobrevivência. O tempo mediano corresponde ao tempo que aproximadamente 50% dos indivíduos esperam sobreviver. Se a curva sofre uma redução abaixo de 0,5, mas não exatamente em 0,5, então o primeiro tempo que apresentar o evento quando a curva cair próxima de 0,5 é usado. Muitos procedimentos computacionais juntam os “degraus” do gráfico com uma linha vertical para melhorar a aparência visual (43,46).

3.4.2 Razão de riscos

Apesar da relação matemática entre a função densidade de probabilidade, a função de sobrevivência $S(t)$ e a função de risco $h(t)$, a natureza condicional da função de riscos faz esta ser conveniente para a comparação do risco de mortalidade entre coortes. Isto ocorre porque a comparação em um dado ponto no tempo baseado na função de riscos envolve somente aqueles indivíduos que sobreviveram àquele ponto. Por contraste, a função de probabilidade e a função de sobrevivência são incondicionais e assim refletem a experiência de falha dos sujeitos que morreram a priori do tempo de interesse, junto com aqueles que sobreviveram. Por esta razão, os modelos de sobrevivência são freqüentemente definidos em termos das funções de risco (43).

Considere duas coortes, uma que é exposta ao fator de risco e outra não exposta. Denote as correspondentes funções de sobrevivência por $S_1(t)$ e $S_2(t)$, e as funções de risco por $h_1(t)$ e $h_2(t)$, respectivamente. A razão das funções de risco

$h_1(t)/h_2(t)$ é crucial para análise de sobrevivência moderna. De particular importância é a situação onde $h_1(t)/h_2(t)$ é constante – isto é, independente de t . Neste caso se escreve

$$HR = \frac{h_1(t)}{h_2(t)},$$

que é chamada de razão de riscos (razão de azares). Quando a suposição acima está assegurada, se diz que $h_1(t)$ e $h_2(t)$ satisfazem à suposição de riscos proporcionais, ou que $h_1(t)$ e $h_2(t)$ são proporcionais. Quando a suposição de riscos proporcionais é satisfeita, o parâmetro HR é uma conveniente medida de efeito para um estudo de coorte, da mesma maneira que a razão de chances (RC) e o risco relativo (RR) servem para outros tipos de estudos. Pela definição de HR, fica claro que similar a RC e RR, a razão de riscos é uma medida multiplicativa de efeito. Por exemplo, se $HR=3$, então, a cada ponto no tempo, um sujeito na coorte exposta tem um risco de experimentar o desfecho que é 3 vezes o risco de um indivíduo que pertence à coorte não exposta. É importante perceber que a suposição de riscos proporcionais especifica que a razão das funções de risco é constante e não a função de risco individual. De fato, imagine que $h_2(t) \geq 0$ tem uma forma funcional arbitrária, para uma dada constante $\psi > 0$, e defina $h_1(t) = \psi \cdot h_2(t)$. Então $h_1(t)$ e $h_2(t)$ são proporcionais e ψ é a razão de riscos. Isto é equivalente a dizer que $S_1(t) = [S_2(t)]^{HR}$. A maioria dos métodos para análise de dados censurados é baseada na suposição de riscos proporcionais. Como resultado, muito da discussão está focada na razão de riscos. Porém, deve ser lembrado que a razão de riscos é uma medida relativa de efeito e, em virtude disso, não informa nada sobre o risco absoluto. Portanto, como parte da análise de sobrevivência é importante examinar as curvas de sobrevivência inteiramente com o objetivo de ganhar uma apreciação mais completa dos riscos de mortalidade (43).

Um passo inicial na análise de um grupo de dados de sobrevivência são medidas numéricas e alguns gráficos dos tempos de sobrevivência para os indivíduos de um grupo particular. Tais medidas resumo podem ser de interesse por si só, ou como precursoras de uma análise de dados mais detalhada. Dados de sobrevida são convenientemente sumarizados através de estimativas da função de sobrevivência e da função de risco. Os métodos para estimativa dessas funções de uma simples

amostra dos dados de sobrevida são ditos não-paramétricos, uma vez que eles não requerem suposições específicas a serem feitas sobre a distribuição latente nos tempos em estudo (42).

Uma vez estimada a função de sobrevida, a mediana e outros percentis da distribuição dos tempos de sobrevivência podem ser estimados. Quando os tempos de sobrevida de dois ou mais grupos estão sendo comparados, uma comparação informal da experiência de sobrevida de cada grupo dos indivíduos pode ser feita usando funções estimadas dos sobreviventes. Porém existem procedimentos mais formais que permite comparar dois grupos de dados de sobrevida. Dois procedimentos não-paramétricos para a comparação de dois ou mais grupos de tempos de sobrevivência conhecidos são os chamados teste de log-rank e teste de Wilcoxon (42).

3.4.3 Comparação de funções de sobrevivência

Freqüentemente existe o interesse em comparar a sobrevivência do grupo dos expostos contra os não expostos. Em adição, a comparação entre curvas fornece informação útil sobre a distribuição de tempo até o evento, que pode ser diferente entre os grupos. Curvas de sobrevivência separadas para diferentes níveis de covariáveis fornecem um efetivo processo de investigação que ajuda a identificar fatores que são influentes em determinadas sobrevivências. As influências dos fatores sobre as curvas de sobrevida devem ser investigadas usando análises multivariadas (40).

A maneira mais simples de comparar tempos de sobrevida obtidos de dois grupos ou mais de indivíduos é criar um único gráfico com as correspondentes estimativas das funções de sobrevida para cada grupo. O resultado do gráfico pode ser bastante informativo (42).

Se uma das curvas estiver sempre acima das demais significa que para qualquer tempo t , a probabilidade estimada de sobrevida além de t é maior para os indivíduos do grupo que apresentar a curva mais “elevada” (42). Embora visualmente as curvas de sobrevivência para dois grupos possam parecer diferentes, é

preciso testar se as curvas de sobrevivência são estatisticamente diferentes através de um teste de hipóteses formal (46).

Existem duas possíveis explicações para se observar uma diferença entre duas funções de sobrevida estimadas. Uma explicação é que existe uma real diferença entre os tempos de sobrevida dos dois grupos de indivíduos, tal que indivíduos de um grupo têm uma experiência de sobrevida diferente daqueles do outro grupo. Uma explicação alternativa é que não existe uma diferença real entre os tempos de sobrevida em cada grupo, e que a diferença que foi observada é meramente o resultado de uma variação ao acaso (42).

Quando não existem dados censurados, testes não-paramétricos padrões podem ser usados para comparar duas distribuições de sobrevivência, como o teste de Mann-Whitney (40). Porém, quando existirem dados censurados o teste de *log-rank* é recomendado. Embora existam outros testes disponíveis, como por exemplo, o teste de Breslow ou de Cox-Mantel, neste trabalho se dará destaque ao teste *log-rank*.

O teste de *log-rank* é provavelmente o teste mais comumente usado para a comparação de distribuições de tempos de sobrevivência (40). A hipótese nula postula que não existe diferença entre as funções de sobrevivência, por exemplo, entre o grupo exposto e o não exposto (46). O teste supõe que as observações são independentes e que a distribuição das censuras é independente da distribuição de sobrevivência. A distribuição da censura pode ser diferente em cada grupo, e o teste pode ser usado para comparar um grupo corrente de tratamento com um controle histórico (46).

Para construir o teste, se começa por considerar separadamente cada tempo de falha em dois grupos de dados de sobrevida. Os grupos serão chamados de Grupo I e Grupo II. Suponha que existam r tempos de falhas distintos, $(t_1 < t_2 < \dots < t_r)$ considerando conjuntamente os dois grupos, e que no tempo t_j , a_{1j} indivíduos no Grupo I e a_{2j} no Grupo II apresentam o evento em estudo (genericamente denotado como falha), para $j = 1, 2, \dots, r$. A menos que dois ou mais indivíduos no grupo tenham o mesmo tempo de falha, os valores de a_{1j} e a_{2j} serão 0 ou 1. Suponha que existam n_{1j} indivíduos em risco de falha no primeiro grupo antes do tempo t_j e que existem n_{2j} indivíduos em risco no segundo grupo.

Conseqüentemente, no tempo t_j , existirão $a_j = a_{1j} + a_{2j}$ falhas no total de $n_j = n_{1j} + n_{2j}$ indivíduos em risco (42).

Uma maneira de avaliar a veracidade da hipótese nula é considerar a extensão da diferença entre o número observado de indivíduos nos dois grupos que apresentaram o evento em cada um dos tempos de falha, e o correspondente número esperado sob a hipótese nula. A informação sobre a extensão destas diferenças pode então ser combinada para cada um dos tempos de falha (42).

Uma tabela 2x2 é formada para cada tempo distinto de falha, onde são considerados os grupos (nas linhas) e o número de falhas e não-falhas naquele período de tempo nas colunas. Os totais marginais (número total de indivíduos em risco depois de t_j , número total de indivíduos que falharam e número total de indivíduos que sobreviveram além de t_j) são considerados fixos e, sob a hipótese nula, as quatro entradas nessa tabela são somente determinadas pelo valor de a_{1j} , o número de falhas no tempo t_j no Grupo I. Sendo assim, os valores do Grupo I serão levados em consideração para calcular a estatística do teste (42).

Os números esperados de falhas no intervalo são determinados assumindo que a_j segue uma distribuição hipergeométrica. Assim, sob hipótese nula, a probabilidade de falha no tempo t_j não depende do grupo ao qual o indivíduo pertence e, então, a probabilidade de falha no tempo t_j será apenas a_j/n_j . Multiplicando essa razão por n_{1j} , tem-se e_{1j} , que é o número esperado de falhas no Grupo I até o tempo t_j (42).

O próximo passo, então, é combinar as informações de cada uma das tabelas criadas para se ter uma medida global de todos os desvios dos valores observados de falha em relação ao seu número esperado. A maneira mais simples de fazer isso é somar as diferenças $(a_{1j} - e_{1j})$, onde e_{1j} é o número esperado de falhas no tempo t_j , para todos os r intervalos de tempo de falha, nos dois grupos. O resultado da estatística é dado por

$$U_L = \sum_{j=1}^r (a_{1j} - e_{1j}) \quad (42).$$

Note que sob a hipótese nula estatística do teste assume valor igual a zero, uma vez que o número esperado de falhas é igual ao número observado. Além disso,

como os tempos de falha são independentes, a variância de U_L simplesmente é a soma das variâncias de a_{1j} , dadas por

$$v_{1j} = \frac{n_{1j}n_{2j}a_j(n_j - a_j)}{n_j^2(n_j - 1)} \quad (42).$$

Então, a variância de U_L é

$$\text{var}(U_L) = V_L = \sum_{j=1}^r v_{1j} \quad (42).$$

Além disso, quando o número de tempos de falhas não é tão pequeno U_L tem distribuição aproximadamente normal e, portanto, seu quadrado tem distribuição aproximada qui-quadrado com um grau de liberdade, ou seja, $U_L^2/V_L \sim \chi_1^2$ (42).

A estatística $W_L = U_L^2/V_L$ resume qual é a extensão dos tempos de sobrevida nos dois grupos de dados a partir dos tempos esperados sobre a hipótese nula de que não existe diferença entre os grupos. Quanto maior o valor da estatística, maior a evidência contra a hipótese nula (42).

Reforçando o que foi citado anteriormente, uma avaliação informal da adequação da suposição de riscos proporcionais pode ser feita a partir de um gráfico das funções de sobrevida para os dois grupos. Se as duas funções de sobrevivência estimadas não se cruzam, a suposição de riscos proporcionais pode ser justificada, e o teste de log-rank é apropriado. As funções de sobrevida estimadas podem se cruzar mesmo se as correspondentes funções de risco verdadeiras são proporcionais, meramente devido ao acaso (42).

A menos que um gráfico das funções de sobrevida estimadas, ou dados prévios, indiquem que existem boas razões para duvidar da suposição de riscos proporcionais, o teste de log-rank deveria ser usado para testar a hipótese de igualdade de duas funções de sobrevida (42).

3.4.4 Modelando dados de sobrevida

Os métodos não-paramétricos, tal como, Kaplan-Meier, podem ser úteis na análise de uma amostra simples de dados, ou na comparação de dois ou mais grupos em relação aos tempos de sobrevida. Porém, na maioria dos estudos que deram origem aos dados de sobrevida, informações suplementares sobre os indivíduos também são registradas, tais como sexo, idade, estilo de vida, etc. Para

incorporar na análise essas variáveis explanatórias são necessários métodos mais apurados (40,42).

Modelando a função de risco

Procedimentos de modelagem de dados de sobrevivência permitem explorar como é a experiência de sobrevida de um grupo de pacientes dependendo dos valores de um ou mais variáveis explanatórias. Como as variáveis adicionais podem ter alguma influência no tempo de sobrevida do indivíduo, é importante levar em consideração estas variáveis quando avaliar a extensão das diferenças entre grupos (42).

Na análise de dados de sobrevida, o interesse central está no risco de falha em qualquer tempo depois do tempo de origem do estudo. Assim, a função de risco é modelada diretamente na análise de sobrevivência. Os modelos resultantes são um pouco diferentes na forma dos modelos lineares encontrados na análise de regressão e na análise de dados de delineamento de experimentos, onde a dependência da resposta média, ou alguma função dela, sobre certas variáveis exploratórias é modelada. Porém, alguns dos princípios e procedimentos de modelos lineares também são empregados na modelagem de dados de sobrevida (42).

Existem duas razões principais para modelar dados de sobrevivência. Um objetivo do processo de modelagem é determinar que combinação de possíveis características afeta a forma da função de risco. Em particular, o efeito que o “tratamento” tem no risco de falha pode ser estudado, e pode também estender a análise para entender como outras variáveis afetam a função de risco. Outra razão para modelar a função de risco é obter uma estimativa da função para um indivíduo, mas também pode ser utilizada para estimar a função de sobrevida. Isto levará a uma estimativa de quantidades tais como o tempo mediano sobrevivência, que será uma função das variáveis explanatórias no modelo. O tempo mediano de sobrevida poderia então ser estimado para pacientes atuais ou futuros com valores particulares para as variáveis explanatórias utilizadas no modelo. A estimativa resultante poderia ser particularmente útil em criar um regime de tratamento, ou em aconselhar o pacientes sobre seu prognóstico (42).

O modelo básico para dados de sobrevivência a ser considerado nesse trabalho é o modelo de riscos proporcionais proposto por Cox (1972). Embora o modelo seja baseado na suposição de riscos proporcionais, nenhuma forma particular de distribuição de probabilidade é assumida para os tempos de sobrevivência, sendo referido como um modelo semi-paramétrico.

Modelo de riscos proporcionais

Na maioria dos estudos, existe o interesse de avaliar o efeito de um ou mais tratamentos ou exposições no desfecho de interesse, sozinho ou ajustado para outras variáveis explanatórias. Na análise de sobrevivência o desfecho de interesse é o tempo de sobrevivência, e um dos objetivos usualmente é comparar os tempos de sobrevivência entre grupos ou avaliar a relação entre a exposição/covariáveis com o tempo de sobrevivência. Métodos padrão de análise de dados, tais como, teste t, regressão logística ou linear, não podem ser aplicados aos dados de sobrevivência porque eles não dão conta da censura. Se as observações censuradas forem excluídas da análise os resultados poderão ser viesados (46).

O modelo será agora generalizado para a situação onde o risco de falha em um tempo particular depende dos valores observados x_1, x_2, \dots, x_p das p variáveis explanatórias. O modelo especifica que a função de risco do i -ésimo sujeito, no tempo t , pode ser escrita como um risco basal $h_0(t)$, multiplicado pela componente sistemática que envolve as variáveis explanatórias, ou seja,

$$h_i(t) = h_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}).$$

O ajuste do modelo de riscos proporcionais para um grupo de dados de sobrevivência observados implica em estimar os coeficientes das variáveis explanatórias $\beta_1, \beta_2, \dots, \beta_p$. Estes dois componentes do modelo podem ser estimados separadamente, sendo que a função de risco base $h_0(t)$ é estimada de forma não paramétrica. Os β 's são estimados primeiro e estas estimativas são então usadas para construir uma estimativa da função de risco base. Este resultado é importante, uma vez que isto significa que para fazer inferências sobre os efeitos das p variáveis

explanatórias no risco relativo $h_i(t)/h_0(t)$, não é necessária uma estimativa de $h_0(t)$ (42).

Suponha que os dados estão disponíveis para n indivíduos, entre os quais existem r tempos distintos de falhas e $n - r$ tempos de sobrevivência censurados a direita. Assuma, por enquanto, que apenas um indivíduo morre em cada tempo de falha, ou seja, que não existem empates nos tempos de sobrevida. Os r tempos de falha ordenados são denotados por $t_{(1)} < t_{(2)} < \dots < t_{(r)}$, tal que $t_{(j)}$ será o j -ésimo tempo ordenado de falha. O grupo de indivíduos que estão sob risco no tempo t_j será denotado por $R(t_j)$, denominado de grupo em risco no tempo t_j , e é composto pelos indivíduos que estão vivos ou não censurados no tempo a priori de t_j . (42).

O grupo de valores das variáveis explanatórias no modelo de riscos proporcionais será representado pelo vetor x , para o qual $x = (x_1, x_2, \dots, x_p)$. A função $h_0(t)$ é interpretada como a função de risco para um indivíduo para o qual os valores de todas as variáveis explanatórias são nulos. Assim, para o i -ésimo indivíduo o modelo pode ser escrito como

$$h_i(t) = \psi(x_i)h_0(t),$$

onde $\psi(x_i)$, é a função dos valores do vetor das variáveis explanatórias. A função $\psi(\cdot)$ pode ser interpretada como o risco de apresentar o desfecho de interesse no tempo t para um indivíduo cujo vetor de variáveis explanatórias é x_i , relativo ao risco para um indivíduo que tem $x=0$ (42).

A função de verossimilhança para o modelo de riscos proporcionais da equação $h_i(t) = h_0(t)\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$ é dado por

$$L(\beta) = \prod_{j=1}^r \frac{\exp(\beta' x_j)}{\sum_{l \in R(t_j)} \exp(\beta' x_j)},$$

em que x_j é o vetor de covariáveis para um indivíduo que morre no j -ésimo tempo ordenado de falha, t_j . O somatório no denominador da função de verossimilhança é a soma dos valores de $\exp(\beta' x)$ levando em conta todos os indivíduos que estão em risco no tempo t_j . Note que o produto é tomado levando em consideração todos os indivíduos que tiveram seus tempos de falhas registrados. Indivíduos para os quais os tempos de sobrevida são censurados não contribuem para o numerador da função log-verossimilhança, mas eles entram no somatório dos grupos de risco em tempos

de falha que ocorrem antes do tempo da censura. Além disso, a função de verossimilhança depende somente do ordenamento dos tempos de falha, uma vez que estas determinam o grupo de risco a cada tempo de falha. Conseqüentemente, inferências sobre os efeitos das variáveis explanatórias na função de risco dependem somente da ordem dos tempos de sobrevivida (42).

Como o risco relativo $\psi(\mathbf{x}_i)$ não pode ser negativo, é conveniente escrevê-lo como $\exp(\eta_i)$, onde η_i é uma combinação linear das p variáveis explanatórias em \mathbf{x}_i . Portanto,

$$\eta_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi},$$

para que $\eta_i = \sum_{j=1}^p \beta_j x_{ji}$. A quantidade η_i é chamada de componente linear do modelo, mas é também conhecida como o escore do risco para o i -ésimo indivíduo. O modelo generalizado de riscos proporcionais

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) h_0(t)$$

também pode ser escrito na forma

$$\log \left\{ \frac{h_i(t)}{h_0(t)} \right\} = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi},$$

ou seja, também pode ser considerado como um modelo linear para o logaritmo da razão de risco (42).

Note que não há um termo constante (intercepto) no componente linear do modelo de riscos proporcionais. Se um termo constante β_0 , digamos, for incluído, a função de risco de base poderia simplesmente ser redimensionada pela divisão de $h_0(t)$ por $\exp(\beta_0)$, e o termo constante seria cancelado (42).

3.4.5 Análise de sobrevivida com truncamento à esquerda

Um tipo de dado incompleto que ocorre em certas situações é a perda de informação devido à exclusão do indivíduo do estudo porque já havia experimentado o evento antes do início do estudo e, assim, não pode ser observado. Por exemplo, em alguns estudos com gestantes as participantes são arroladas em clínicas de pré-natal em idades gestacionais diferentes. Nessa situação pode ocorrer truncamento à esquerda, pois o tempo zero (a concepção) não é observado numa cronologia relevante. Calcular os efeitos estimados sem levar em conta o truncamento à

esquerda pode introduzir vieses porque a entrada no estudo é diferenciada pelo desfecho e pode ser também diferenciado pelo *status* de exposição (12). Este tipo de truncamento é comum na análise de dados médicos ou epidemiológicos e, se for ignorado, os resultados podem ser gravemente viesados (47).

É possível considerar o tempo de observação com truncamento à esquerda na análise de sobrevivência através de uma série de conjuntos de risco que incluem somente aquelas mulheres que entraram no estudo a tempo de ser observada a perda da gravidez num determinado instante. A suposição é que as mulheres que tiveram adicionada a observação para uma dada idade gestacional são representativas de todas as gravidezes sobreviventes para aquela idade gestacional na população fonte (incluindo aquelas que nunca estiveram sob observação para aquela idade gestacional e aquelas que já estão sob observação). Na análise de regressão de Cox que acomoda o truncamento à esquerda as gravidezes contribuem para a probabilidade condicional em terem entrado em observação naquele tempo, em oposição à regressão de Cox tradicional, onde todas as gravidezes são tratadas como se todas contribuíssem para a probabilidade desde o início do estudo (12).

Além disso, é importante ressaltar que como as mulheres são recrutadas após a concepção, os grupos de risco são baseados nas idades gestacionais e possíveis erros de mensuração na estimação da idade gestacional podem introduzir vieses. Mais especificamente, se a idade gestacional estimada em uma dada gestação é incorreta, a gravidez poderia ser inapropriadamente incluída ou excluída de um dos conjuntos de risco. Conforme relatado por Howard et al. (48) existem numerosas fontes potenciais de erro de estimação da idade gestacional para uma gestante. Como a idade gestacional é tipicamente calculada a partir do primeiro dia do último período menstrual da mulher, a idéia é que todas as concepções ocorrem no mesmo dia, classicamente considerado o 14º dia do ciclo menstrual. A data é primeiramente fornecida pela própria mulher, sendo possível ser revisada por base de um julgamento físico ou de um ultra-som. Auto-resposta mostra uma preferência por dígitos, com dias do mês que são múltiplos de 5 sendo particularmente super-representados. Além disso, algumas mulheres são incapazes de lembrar a data do último período menstrual, sendo esta mais esquecida entre mulheres de classes sociais baixas e por grávidas com desfechos adversos. Mesmo se a mulher informar a

data corretamente, sua idade gestacional ainda pode ser estimada de maneira inacurada, porque existem mulheres que apresentam ciclos menstruais maiores. Assim, mulheres com a mesma estimativa de idade gestacional podem não ter ficado grávidas pela mesma quantidade de tempo, porque elas poderiam ter concebido em dias diferentes.

De uma maneira geral, a literatura sugere que o exame de ultra-som prediz o parto com mais confiança que a auto-resposta das gestantes, mas nenhum método é perfeito (48). Sendo assim, e devido à maioria dos estudos trabalharem com gestantes em observação após a concepção, é importante adotar uma análise estatística que diminua o impacto dos vieses, e que eles sejam em menor número possível. Então, como discutido por Howards et al. (12) em outro artigo, se o truncamento à esquerda é considerado um diferencial através do tempo de exposição, a regressão logística pode produzir uma estimativa de razão de chances (RC) viesada porque o tempo em risco observado para um determinado grupo dos expostos é maior do que para outro grupo também exposto, como os censurados, por exemplo. Em tal situação, os modelos de sobrevivência seriam mais corretos do que os modelos logísticos porque levam em conta o tempo que cada gravidez está sob observação. Vale lembrar que a análise estatística informa uma razão de risco (HR), e não razão de chances (RC) ou risco relativo (RR).

O modelo de regressão de Cox com truncamento à esquerda

O modelo de Cox escolhido para a análise dos dados pode ser de dois tipos: paramétrico ou semi-paramétrico. A diferença é que o modelo paramétrico assume que a função densidade de probabilidade do tempo para um particular evento segue uma distribuição específica, tal como a distribuição exponencial, enquanto que o modelo semi-paramétrico não assume uma forma funcional (47).

Partindo do modelo básico, admita que T representa o tempo até a ocorrência de algum evento específico. A variável aleatória T é não-negativa e proveniente de uma população homogênea. É conveniente lembrar as quatro funções que caracterizam a distribuição de T , que são: a *função de sobrevivência* $S(t)$; a *função de risco* $h(t)$; a *função densidade de probabilidade* $f(t)$; e o *tempo médio residual*. Se qualquer uma dessas quatro funções é conhecida, então as outras três

podem ser determinadas. Na prática, essas quatro funções costumam estar aliadas com outra quantidade útil, a função de risco acumulado (44).

Vários tipos de curvas de sobrevivência podem ser construídos, mas o ponto mais importante que deve ser notado é que todas têm as mesmas propriedades básicas. Elas são monótonas, funções não crescentes, assumem o valor igual a 1 no tempo zero e tendem para zero conforme o tempo se aproxima do infinito. Sua taxa de declínio, é claro, varia de acordo com o risco de experimentar o evento no tempo t , mas é difícil determinar a essência de um padrão de falha simplesmente olhando para a curva de sobrevivência. Ainda, as curvas podem ser muito úteis quando se deseja comparar dois ou mais padrões de falha (44).

A função de risco é fundamental na análise de sobrevivência e na epidemiologia é conhecida como taxa de falha em uma idade específica, ou simplesmente, razão de risco. Esta função é particularmente útil para determinar a distribuição de falha apropriada utilizando informação qualitativa sobre o mecanismo de falha e para descrever a maneira que a chance de experimentar o evento muda com o tempo. Existem muitas formas para a razão de risco e a única restrição é que ela não seja negativa, isto é, $h(t) \geq 0$ (44).

De fato, a função de risco é geralmente mais informativa sobre o mecanismo de falha envolvido no estudo que a função de sobrevivência. Por esta razão, a função de risco pode ser um método dominante para resumir os dados de sobrevivência (44).

A função de vida média residual ou vida mediana informa para um indivíduo na idade x o tempo de “vida” restante. Pode ser mostrado ainda que, o tempo médio residual é a área abaixo da curva de sobrevivência à direita de t , dividida por $S(t)$.

Até este ponto, foi tratada exclusivamente a idéia de quando o tempo T pertence a uma população homogênea. Porém, um problema frequentemente encontrado na análise dos dados de sobrevivência é o ajustamento de função de sobrevivência levando em consideração a representação de informações de variáveis explanatórias e tais populações apresentam certa heterogeneidade (44).

Imagine então a variável $T > 0$ e um vetor de variáveis explanatórias $X = (X_1, \dots, X_k)$ associadas ao tempo de falha T . O vetor \mathbf{X} pode incluir variáveis

quantitativas, qualitativas e/ou tempo dependentes. Frequentemente a questão de maior interesse é acerca da relação entre o tempo de falha T e uma ou mais variáveis explanatórias (44).

Em estudos de sobrevivência a censura e o truncamento precisam ser cuidadosamente considerados quando é construída a função de verossimilhança. Uma observação correspondendo a um tempo exato do evento fornece informações da probabilidade de que o evento ocorra naquele tempo que é aproximadamente igual à função de densidade de T naquele tempo. Para dados truncados à esquerda, estas probabilidades são substituídas pelas probabilidades condicionais apropriadas. Mais especificamente, a verossimilhança para estes dados deve incorporar o componente $f(t)/S(Y)$, onde $f(t)$ é a função densidade de probabilidade (44).

Para derivar os estimadores da função de sobrevivência para dados truncados à esquerda é necessário, primeiro, rever alguns conceitos. Para dados que são apenas censurados à direita define-se que $t_1 < t_2 < \dots < t_D$ como os tempos distintos de falha, L_i como sendo, no caso, a idade gestacional aleatória em que uma gestante foi incorporada ao estudo, T_j o tempo que a gestante manifestou o evento em estudo ou a censura e a_i como o número de indivíduos que experimentaram o evento de interesse no tempo t_i . Permanece sendo necessário o número de indivíduos que estão em risco de manifestar o evento de interesse no tempo t_i , chamado Y_i . Para dados censurados à direita, esta quantidade é o número de indivíduos no estudo no tempo 0 com o estudo estando, pelo menos, no tempo t_i . Para dados truncados à esquerda, Y_i é redefinido como o número de indivíduos que entraram no estudo a priori do tempo t_i e que têm um tempo de estudo de pelo menos t_i , ou seja, Y_i é o número de indivíduos que estão $L_i \leq t_i \leq T_j$ (44).

Definindo Y_i dessa maneira para dados truncados à esquerda, todos os procedimentos de estimação definidos anteriormente são aplicáveis nesse momento. Porém, deve-se tomar cuidado ao interpretar as estatísticas geradas, pois, o estimador produto-limite de Kaplan-Meier da função $S(t)$ no tempo t é agora um estimador de probabilidade de sobreviver além de t , condicionado a sobreviver ao menor tempo de entrada L , ou seja, $\Pr[T > t | T \geq L] = S(t) / S(L)$.

Alguns cuidados ao aplicar este estimador são necessários. Para dados truncados à esquerda, é possível que o número de indivíduos em risco seja bastante

pequeno para valores pequenos de t_i . Se, para algum valor de t_i , Y_i for igual a a_i , então, o estimador do produto-limite será zero para todos os valores de t além desse ponto. Nestes casos, é comum estimar a função de sobrevivência condicional a sobreviver um valor de tempo onde isto não acontece considerando assim somente aqueles tempos de falhas além desse ponto (44).

Sendo assim, o estimador padrão para a função de sobrevivência, proposta por Kaplan-Meier, chamada de estimador Produto-Limite é dado por

$$\hat{S}_L(t) = \begin{cases} 1 & \text{se } t < t_i, \\ \prod_{L \leq t_i \leq t} \left[1 - \frac{a_i}{Y_i} \right] & \text{se } t \geq L \end{cases}$$

O estimador Produto-Limite é uma função escada com pulos (saltos) nos tempos em que são observados os eventos. O tamanho dos pulos depende não só do número de eventos observados para cada tempo t_i , mas também do padrão das observações truncadas antes de t_i (44).

Logo, para formular o modelo de riscos proporcionais para um grupo de covariáveis X , modela-se a razão de risco condicional de T , dado X e $T > L$, isto é,

$$h(t | X, T > L) \cong \frac{P(T = t | X, T > L)}{P(T \geq t | X, T > L)}$$

4. OBJETIVO

Estimar o risco de anomalia congênita para a exposição ao misoprostol durante a gestação através do modelo de Cox com truncamento à esquerda no Estudo Brasileiro do Diabetes Gestacional (EBDG), comparando estas estimativas com os resultados gerados pelo modelo de regressão logística e com o modelo de Cox com e sem truncamento à esquerda.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- (1) Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic Measures of the Course and Outcome of Pregnancy. *Epidemiologic Reviews* 2002; 24(2):91-101
- (2) Koren G, Pastuszak A, Shinya I. Drugs and Pregnancy. *New England Journal of Medicine* April 1998; 338(16):1128-1137.
- (3) Dal Pizzol TS. Risco e benefícios para o feto e recém-nascido de medicamentos utilizados na gestação: misoprostol e antianêmicos. Tese de Doutorado pela Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Epidemiologia. Porto Alegre, BR-RS, 2006.
- (4) Victora CD, Barros FC. Infant mortality due to perinatal causes in Brazil: trends, regional patterns and possible interventions. *São Paulo Medical Journal* 2001; 119(1):33-42.
- (5) Mengue SS, Schenkel, Duncan BB, Schmidt MI. Prevalence and Clinical Correlates of Unsuccessful Use of Drugs to Induce Menstruation. *Contraception* 1998; 57:93-97.
- (6) Costa SH, Vessey. Misoprostol and illegal abortion in Rio de Janeiro, Brazil. *Lancet* 1993; 341:1258-1261.
- (7) Coêlho HLL, Teixeira AC, Santos AP, Forte EB, Morais SM, La Vecchia C et al. Misoprostol and illegal abortion in Fortaleza, Brazil. *Lancet* 1993; 341:1261:1263.
- (8) Goldberg AB, Greenber BS, Darney PD. Misoprostol and pregnancy. *New England Journal of Medicine* 2001; 344:38-46.
- (9) Mengue SS, Dal Pizzol TS. Misoprostol, aborto e malformações congênitas. *Brás Ginecol Obstet.* 2008; 30(6):271-273.
- (10) Fonseca W, Misago C, Correia LL, Parente JAM, Oliveira FC. Determinantes do aborto provocado entre mulheres admitidas em hospitais em localidades da região Nordeste do Brasil. *Revista de Saúde Pública* 1996; 30(1):279-286.
- (11) Fonseca W, Misago C, Freitas P, Santos E, Fernandes L, Correia L. Características sócio-demográficas reprodutivas e médicas de mulheres admitidas por aborto em hospital da Região Sul. *Cadernos de Saúde Pública* 1998; 14(2):279-286
- (12) Howards PP, Hertz-Picciotto I, Poole C. Conditions for Bias from Differential Left Truncation. *American Journal of Epidemiology* 2007; 165(4):444-452.
- (13) Loebstein R, Koren G. Clinical relevance of therapeutic drug monitoring during pregnancy. *Therapeutic Drug Monitoring.* 2002; 24(1):15-22

- (14) Vallance P. Drugs and the fetus. *BMJ* 1996; 312:1053-1054.
- (15) Czeizel AE. Recommendation to avoid all drugs during first trimester is unrealistic. *BMJ* 1996; 313:424-425.
- (16) Mengue SS, Schenkel EP, Duncan BB, Schmidt MI. Uso de medicamentos por gestantes em seis cidades brasileiras. *Revista de Saúde Pública* 2001; 35(5):415-420.
- (17) Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Ações Programáticas Estratégicas. Pré-natal e puerpério. Atenção qualificada e humanizada. Manual técnico. Brasília: Ministério da Saúde, 2005: 67-68.
- (18) Schönhofer PS. Misuse of misoprostol as an abortifacient may induce malformations. *Lancet* 1991; 337(1534):1535.
- (19) Song J. Use of misoprostol in obstetrics and gynecology. *Obstetrical and Gynecological Survey*, August 2000; 55(8):503-510.
- (20) Coêlho HLL. Misoprostol: a solução não é tão simples. *Revista de Saúde Pública*, 1998; 32(4):394-395.
- (21) Misago C, Fonseca W, Correia LL, Fernandes LM, Campbell O. Determinants of abortion among women admitted to hospitals in Fortaleza, North Eastern Brazil. *International Journal of Epidemiology* 1998; 27:833-839.
- (22) Gallo MF, Cahill S, Castleman L, Mitchell MH. A systematic review of more than one dose of misoprostol after mifepristone for abortion up to 10 weeks of gestation. *Contraception* 2006; 74:36-41.
- (23) Marions L. Mifepristone dose in the regimen with misoprostol for medical abortion. *Contraception* 2006; 74:21-25.
- (24) Fontenelle L, Araujo APQC, Fontana RS. Síndrome de Moebius. *Arquivos de Neuropsiquiatria* 2001; 59(3-B):812-814.
- (25) Schüller L, Pastuszak AL, Sanseverino MTV, Orioli IM, Brunoni D, Ashton-Prolla P et al. Pregnancy outcome after exposure to misoprostol in Brazil: a prospective, controlled study. *Reproductive Toxicology* 1999; 13(2):147-151.
- (26) Horovitz DDG, Llerena Jr JC, Mattos RA. Atenção aos defeitos congênitos no Brasil: panorama atual. *Cadernos de Saúde Pública* 2005; 21(4): 1055-1064.
- (27) Castro MLS, Cunha CJ, Moreira PB, Fernández RR, Garcias GL, Martino-Roth MG. Frequência das malformações múltiplas em recém-nascidos na Cidade de Pelotas, Rio Grande do Sul, Brasil, e fatores sócio-demográficos associados. *Cadernos de Saúde Pública* 2006; 22(5): 1009-1015.

- (28) Basso O, Olsen J, Christensen K. Recurrence Risk of Congenital Anomalies – the Impact of Paternal, Social, and Environmental Factors: A Population-based Study in Denmark. *American Journal of Epidemiology* 1999; 150(6): 598-604.
- (29) Morales-Suárez-Varela MM, Bille C, Christensen K, Olsen J. Smoking Habits, Nicotine Use, and Congenital Malformations. *American College of Obstetricians and Gynecologists* 2006; 107(1): 51-57.
- (30) Wolfe H. High Prepregnancy Body-Mass Index – A Maternal-Fetal Risk Factor. *The New England Journal of Medicine* 1998; 338(3): 191-192.
- (31) Anderson JL, Waller DK, Confield MA, Shaw GM, Watkins ML, Werler MM. Maternal Obesity, Gestational Diabetes, and Central Nervous System Birth Defects. *Epidemiology* 2005; 16:87-92.
- (32) Shaw GM, Todoroff K, Schaffer DM, Selvin S. Maternal height and prepregnancy body mass index as risk factors for selected congenital anomalies. *Pediatric and Perinatal Epidemiology* 2000; 14:234-239.
- (33) Aguiar MJB, Campos AS, Aguiar RALP, Lana ANA, Magalhães RL, Babeto LT. Defeitos de fechamento do tubo neural e fatores associados em recém-nascidos vivos e natimortos. *Jornal de Pediatria* 2003; 79(2):129-134.
- (34) Ray JG, O'Brien TE, Chan WS. Preconception care and the risk of congenital anomalies in the offspring of women with diabetes mellitus: a meta-analysis. *Q J Med* 2001; 94:435-444.
- (35) Queiber-Luft A, Stolz G, Wiesel A, Schlaefter K, Spranger J. Malformations in newborn: results based on 30940 infants and fetuses from the Mainz congenital birth defect monitoring system (1990-1998). *Arch Gynecol Obstet* 2002; 266:163-167.
- (36) Watkins ML, Rasmussen SA, Honein MA, Botto LD, Moore CA. Maternal Obesity and Risk for Birth Defects. *Pediatrics* 2003; 111:1152-1158.
- (37) Rasmussen SA, Moore CA, Paulozzi LJ, Rhodenbiser EP. Risk for birth defects among premature infants: A population-based study. *The Journal of Pediatrics* 2001; 138(5): 668-673.
- (38) Kramer MS. The Epidemiology of Adverse Pregnancy Outcomes: An Overview. *The Journal of Nutrition* 2003; Supplement:1592S-1596S.
- (39) Parada CMGL, Pelá NTR. Idade Materna como Fator de Risco: Estudo com Primigestas na Faixa Etária Igual ou Superior a 28 anos. *Revista latino-americana de enfermagem* 1999; 7(4):57-64.
- (40) Stevenson M. An introduction to Survival Analysis. Study Guide, Massey University, Palmerston North, New Zealand 2007; 1-31.

- (41) Oakes D. Biometrika Centenary: Survival analysis. *Biometrika* 2001; 88(1):99-142.
- (42) Collet D. Modelling Survival data in medical research. London, United Kingdom: Chapman & Hall, 1994.
- (43) Newman SC. Biostatistical Methods in Epidemiology. Wiley, 2001.
- (44) Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data. New York, Springer-Verlag 1996.
- (45) Guang G. Event-History Analysis for Left-Truncated Data. *Sociological Methodology*, Peter V. Marsden; 1993:217-243.
- (46) Rao SR, Schoenfeld DA. Survival Methods. *Circulation* 2007; 115:109-113.
- (47) Pelz CJ, Klein JP. Analysis of Survival Data: A comparison of three major statistical packages (SAS, SPSS, BMDP). Medical College of Wisconsin 2003.
- (48) Howards PP, Hertz-Picciotto I, Weinberg CR, Poole C. Misclassification of Gestational Age in the Study of Spontaneous Abortion. *American Journal of Epidemiology* 2006; 164:1126-1136.

6. ARTIGO

Risco de anomalia congênita por misoprostol através do modelo de Cox com truncamento à esquerda

Marília Canabarro Zordan¹
Álvaro Vigo^{1,3}
Tatiane da Silva Dal Pizzol^{1,2}
Sotero S. Mengue¹
e Colaboradores

¹ Programa de Pós-graduação em Epidemiologia, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul

² Departamento de Produção e Controle de Medicamentos, Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul,

³ Departamento de Estatística, Instituto de Matemática, Universidade Federal do Rio Grande do Sul

Correspondência:

Marília Canabarro Zordan
Faculdade de Medicina. UFRGS
Ramiro Barcelos, 2600, 4º andar.
Porto Alegre, RS. CEP 90035-003.
e-mail: mariliazordan@yahoo.com.br

Resumo:

Introdução: Na maioria de estudos com gestantes a entrada no estudo ocorre em idades gestacionais diferentes. O tempo entre a concepção e o evento de interesse usualmente é truncado à esquerda e não considerar este aspecto na análise pode produzir vieses. Neste estudo é investigada a associação entre anomalia congênita e o uso do misoprostol, comparando três métodos de análise. **Métodos:** Coorte de 4251 mulheres com 20 anos de idade ou mais, arroladas entre a 21^a e a 28^a semana de gestação e acompanhadas até o parto, procedentes de serviços de pré-natal do SUS localizados em seis capitais brasileiras. Foram comparadas a razão de azares produzida pela regressão de Cox com e sem truncamento à esquerda, e a razão de chances estimada pela regressão logística. **Resultados:** O uso de misoprostol foi identificado como fator de risco para anomalia congênita pelo modelo logístico (RC=2,9; IC95%: 1,12–7,44) e pelos modelos de Cox sem truncamento à esquerda (HR=3,3; IC95%:1,28-8,43) e com truncamento à esquerda (HR=4,8; IC95%:1,69-13,79), porém os preditores dos modelos multivariados variaram. **Conclusão:** A escolha do método de análise influencia os modelos finais e a magnitude da associação estimada. A característica de temporalidade incluída no modelo de Cox com truncamento à esquerda agrega informações importantes por se tratar de um desfecho que altera suas probabilidades de acontecimento com o desenvolver da gestação.

Introdução

Em uma reprodução bem sucedida, a variedade de eventos que pode acontecer entre a concepção e o nascimento de uma criança saudável, é biológica e epidemiologicamente complexa (1). A necessidade de uma atenção cuidadosa ao tempo dos eventos na epidemiologia reprodutiva aplica-se ao delineamento dos estudos e à análise dos dados. É necessário o reconhecimento explícito do tempo de ocorrência do desfecho, da exposição, e o tempo em que a gravidez passou a ser conhecida.

No que diz respeito à farmacoterapia, para a maioria dos medicamentos existem poucas informações sobre os efeitos no feto no momento em que o produto é comercializado e as existentes geralmente provêm de estudos em animais. Embora se tenha percebido que todo medicamento que causa teratogenicidade em humanos também tem causado efeitos similares em animais, alguns ainda são exceção, como é o caso do misoprostol (2,3). Considerando que, por razões éticas, as gestantes são excluídas da maioria dos ensaios clínicos, a avaliação da teratogenicidade dos medicamentos nos estudos pré-comercialização é limitada. Estudos epidemiológicos, tais como estudos de coorte e caso-controle, têm sido uma alternativa importante também para avaliação do risco teratogênico (3).

O misoprostol foi introduzido no Brasil em 1986 para o tratamento de úlcera gástrica e duodenal, mas logo ganhou popularidade como abortivo (4,5). O misoprostol foi comercializado sem restrições nos balcões das farmácias e drogarias até julho de 1991, quando Ministério da Saúde proibiu a venda sem apresentação e retenção da prescrição médica.

Mesmo que diversos estudos tenham sugerido risco associado ao uso desse medicamento e das restrições sanitárias estabelecidas para comercialização no Brasil, estima-se que este fármaco continue sendo utilizado, provavelmente para causar aborto, principalmente em países onde esta prática é proibida, como é o caso do Brasil (3,7).

O risco à saúde materna não é o único problema associado ao uso indiscriminado de substâncias abortivas. Ensaios clínicos controlados demonstram que a falha abortiva com o uso do misoprostol em abortos médicos pode ultrapassar 10%, exigindo emprego de método abortivo adicional para a consumação do

procedimento (6). Outros estudos que consideraram o uso do misoprostol associado a mifepristona encontraram taxas de falhas aproximadamente entre 1% e 10% (8,9). Se sob condições controladas proporcionadas pelos ensaios clínicos, o uso de misoprostol apresenta taxas de falha abortiva indesejáveis, nas situações em que o aborto ilegal é comumente realizado, sem assistência médica adequada e com emprego de doses inadequadas, o misoprostol pode não interromper a gestação, gerando ansiedade quanto ao risco para o feto (3).

Em relação ao tipo de análise estatística utilizada, muitos estudos sobre aborto utilizam regressão logística para analisar a relação entre o desfecho e os fatores de exposição e estimar as medidas de associação. Porém, este método não leva em consideração que as mulheres entram no estudo em tempos diferentes, uma vez que, em geral, são arroladas a partir de clínicas de pré-natal, após a concepção. Sendo assim, as mulheres entram no estudo em diferentes idades gestacionais e, mais importante, uma parte da população fonte é desconhecida por causa das gravidezes perdidas antes do arrolamento. Assim, as mulheres que levaram a gestação até o nascimento têm maior oportunidade de entrar no estudo. Para este tipo de dado, a análise de sobrevivência é recomendada por ser capaz de representar o tempo em observação e, dessa forma, tratar o truncamento à esquerda, através de uma série de conjuntos de risco que incluem somente aquelas mulheres que entraram no estudo a tempo de ser observado o desfecho. A suposição é que as mulheres que tiveram adicionada a observação para uma dada idade gestacional são representativas de todas as gravidezes sobreviventes para aquela idade gestacional na população fonte, incluindo aquelas que nunca estiveram sob observação para aquela idade gestacional e aquelas que já estão sob observação (10).

Dentro deste contexto, o presente trabalho compara empiricamente a associação entre uso de misoprostol e anomalia congênita, utilizando os modelos de regressão logística e regressão de Cox que incorpora e não incorpora truncamento à esquerda.

Material e Métodos

Foram utilizados os dados do Estudo Brasileiro do Diabetes Gestacional (EBDG), uma coorte multicêntrica cujo objetivo foi investigar o diabetes e a intolerância à glicose em gestantes, em relação à prevalência, fatores de risco,

incidência de complicações obstétricas e neonatais, fatores prognósticos e critérios diagnósticos. A primeira fase do estudo incluiu o arrolamento das gestantes, realização da entrevista, tomada de medidas antropométricas e de um teste de tolerância à glicose. Na segunda fase, foi realizado o acompanhamento da gestação, através da revisão de prontuários. A terceira fase foi composta pelas informações sobre o parto e o recém-nascido (3).

Para este trabalho foram utilizados os dados referentes ao uso do misoprostol para induzir a menstruação durante o período gestacional, levantados na primeira fase do EBDG. A amostra foi composta por mulheres de 20 anos de idade ou mais, que se encontravam entre a 21^a e a 28^a semana de gestação, e que foram atendidas em serviços de pré-natal de unidades de saúde do Sistema Único de Saúde (SUS) localizadas em Fortaleza (CE), Manaus (AM), Porto Alegre (RS), Rio de Janeiro (RJ), Salvador (BA) e São Paulo (SP), entre fevereiro de 1991 e junho de 1995, totalizando uma amostra inicial de 5564 gestantes (3).

Foram excluídas 871 gestantes sem informações sobre anomalia congênita ou sobre o uso de misoprostol e, adicionalmente, 442 gestantes do centro de Manaus por apresentar apenas um caso de anomalia congênita, restando 4251 para a análise. Porém, cabe ressaltar que o número de gestantes em cada modelo variou devido à eventual falta de informações sobre os demais preditores.

A entrevista foi estruturada para responder questões relacionadas com aspectos sócio-demográficos, além das questões sobre o uso de tabaco, álcool, drogas e história reprodutiva. Variáveis relacionadas com a gestação em curso, o parto e condições de saúde do neonato foram coletadas do prontuário da mãe. Para investigar o uso específico de produtos para induzir a menstruação, foi perguntado: “Para descobrir se está grávida, a Senhora utilizou algum tipo de remédio para vir a menstruação?” (3).

As informações referentes ao diagnóstico de anomalia congênita foram coletados no prontuário da mãe, (3).

O principal preditor é a exposição ao misoprostol e os outros preditores são idade da mãe, altura da mãe, cor de pele, escolaridade da mãe, índice de massa corporal pré-gravídico, idade gestacional do recém nascido, peso do recém nascido, centro onde foi realizado o atendimento, distúrbios hipertensivos, número de

consultas de pré-natal e aborto prévio. A categorização da maioria dessas variáveis deu-se através de informações retiradas da literatura (12,13,14,15,16).

Para o modelo de regressão de Cox foi considerado como desfecho o número de dias que a gestante ficou em observação, ou seja, o tempo decorrido entre a primeira consulta de pré-natal e o parto, medido pela diferença entre as idades gestacionais de cada momento.

As associações foram estimadas utilizando os modelos de regressão logística, regressão de Cox e regressão de Cox com truncamento à esquerda univariados e multivariados. As variáveis com $p < 0,2$ pelo teste da razão de verossimilhança foram incluídas nos modelos multivariáveis, excluindo uma a uma aquelas com P superior a 5% até chegar ao modelo final.

A diferença do modelo de regressão de Cox com truncamento à esquerda em relação ao tradicional está na contribuição de cada gestante para a verossimilhança condicional. Ao contrário do modelo sem truncamento à esquerda, no qual todas as gestações contribuem para a verossimilhança desde o início do estudo, no modelo que considera o truncamento as gestantes são adicionadas a grupos de riscos específicos, de acordo com o tempo individual de entrada. Por exemplo, uma mulher que entra no estudo no 100º dia de gestação entrará para o grupo de risco que começa no 100º dia, e não participará do grupo de risco que começa no 50º dia de gestação. A variável selecionada para indicar os diferentes tempos gestacionais de entrada das grávidas no estudo foi “idade gestacional na primeira consulta de pré-natal”. A suposição é que as mulheres que tiveram adicionada a observação para uma dada idade gestacional são representativas de todas as gestações sobreviventes para aquela idade gestacional na população fonte (10). Sendo assim, a idade gestacional da primeira consulta é mais informativa que a idade gestacional do arrolamento, uma vez que esse ocorreu apenas entre a 21ª e a 28ª semana de gestação.

As análises foram realizadas com os procedimentos PROC LOGISTIC e PROC PHREG do programa SAS versão 9.1.3.

Resultados

A análise estatística partiu das informações de 4251 mulheres (84,3%) das 5564 arroladas, para as quais está disponível a informação sobre anomalia congênita

e sobre o uso ou não do misoprostol. Foram observados 85 (2,0%) casos de anomalia congênita e 101 (2,4%) gestantes referiram o uso misoprostol. Dentre estas foram observados 5 (4,9%) casos de anomalia, contra 1,9% nas gestantes que não referiram uso de misoprosol. A Tabela 1 descreve as características clínicas e sócio-demográficas da amostra utilizada na análise.

A Tabela 2 resume os resultados dos modelos univariáveis e multivariáveis de regressão logística e Cox com e sem truncamento à esquerda. Os preditores escolaridade da mãe, altura da mãe e história de aborto prévio apresentaram $p > 0,20$ nas análises univariáveis e por isso não foram incluídos em nenhum dos modelos multivariáveis. O uso de misoprostol foi identificado como fator de risco para anomalia congênita pela regressão logística (RC=2,9; IC 95%:1,12–7,44) ajustado por centro e peso do recém-nascido. Utilizando o modelo de Cox sem considerar truncamento à esquerda a estimativa desta associação foi HR=3,3 (IC95%:1,28-8,43), mesmo ajustando por centro, cor de pele, número de consultas de pré-natal, idade gestacional do recém-nascido e peso do recém-nascido. Porém, incorporando truncamento à esquerda, a associação aumenta para HR=4,8 (IC95%:1,69–13,79), no modelo ajustado por idade da mãe, número de consultas de pré-natal, idade gestacional do recém-nascido e peso do recém-nascido. O centro do estudo foi excluído desse modelo por apresentar instabilidade numérica para o Rio de Janeiro. No entanto, no modelo ajustado adicionalmente por centro (sem informações sobre o Rio de Janeiro) e cor de pele a associação não mudou substancialmente (HR=4,1; IC95%:1,39-12,24) (dados não mostrados).

Discussão

Este estudo comparou três estratégias de análise estatística: regressão logística, regressão de Cox sem truncamento à esquerda e regressão de Cox com truncamento à esquerda. A comparação desses três modelos evidenciou que a escolha da análise pode impactar nos resultados obtidos, na magnitude das associações e nas variáveis que compõem o modelo. A magnitude da associação entre anomalia congênita e o uso de misoprostol foi sensivelmente diferente, particularmente para o modelo de Cox com truncamento à esquerda.

Independentemente do modelo utilizado, o uso de misoprostol foi detectado como potencial fator de risco para anomalia congênita. Para todos os preditores, a razão de chances estimada foi, em geral, menor do que a correspondente estimativa de razão de azares (HR), sugerindo que pode existir uma subestimação da associação com o modelo logístico.

Uma característica dos dados deste estudo, já ressaltada por Howards et al, é o fato das gestantes serem arroladas em épocas diferentes, e o tempo, no que diz respeito à gravidez é fundamental, pois vários eventos durante a gestação podem ser dependentes da idade gestacional que a mulher se encontra (13,17). A regressão logística é incapaz de acomodar essas diferenças de tempo e, conseqüentemente, não deveria ser utilizada dada a possibilidade de vieses. Similarmente ao modelo logístico, no modelo de regressão de Cox as gestantes são vistas como tendo sido arroladas no mesmo ponto do tempo. Porém, o modelo de regressão de Cox com truncamento à esquerda consegue lidar com essas entradas em tempos diferenciados.

De um modo geral, os resultados das análises de regressão de Cox com o truncamento à esquerda parecem os mais coerentes. A estimativa de efeito significativa encontrada entre a ingestão do medicamento e o desfecho de anomalia congênita no modelo com truncamento à esquerda pode ser a mais acurada se comparada com as demais análises, apesar da pouca precisão do intervalo de confiança decorrente do pequeno número de casos observados na amostra. As mulheres não são acompanhadas desde o momento da concepção e muitas delas só realizaram sua primeira consulta de pré-natal depois de 12 semanas, após o tempo ideal para iniciar os cuidados com a gravidez. A principal influência que o arrolamento de grávidas em clínicas de pré-natal pode trazer é a perda de informação do que acontece no período entre a concepção e o arrolamento, e não é correto considerar que todas tenham sido arroladas no mesmo momento da gestação.

Embora o misoprostol tenha taxas de falhas consideradas significativas, a sua utilização pode ter possibilitado que muitas mulheres abortassem, e sendo assim, nem ao menos tiveram a possibilidade de serem arroladas. Essa relação foi demonstrada por Schüller et al em um estudo sobre a associação entre aborto e exposição ao misoprostol (18). Ainda, é importante salientar que uma vez que o misoprostol é proibido e seu uso envolve questões morais e religiosas, a acurácia das

respostas sobre a ingestão do medicamento pode estar influenciando na medida de associação, isto é, o resultado pode estar subestimado (3).

Ao contrário da regressão logística, o modelo de Cox com o truncamento à esquerda identificou associação entre anomalia congênita e número de consultas pré-natal, idade da mãe e idade gestacional do recém nascido, que são alguns preditores relatados na literatura; a associação com baixo peso de nascimento foi identificada em ambos os modelos (15,16,20). Porém, a ausência de outros estudos que tenham utilizado o modelo de regressão de Cox com truncamento a esquerda impossibilita comparar as associações estimadas.

É importante mencionar, também, as estimativas da associação mais elevadas para o centro Rio de Janeiro. Isto pode ser explicado pelo fato de que o hospital onde foi conduzido o estudo é especializado no atendimento de gestações com complicações e, portanto, os resultados não podem ser extrapolados para os demais hospitais nesse município.

A utilização de dados secundários pode constituir um potencial limitador deste trabalho pela ausência de algumas informações sobre o período da gestação em que o medicamento foi utilizado, o desfecho anomalia congênita e a respeito da idade gestacional na primeira consulta de pré-natal, utilizada na especificação do truncamento à esquerda. O baixo número de mulheres que referiram o uso de misoprostol em alguns centros e a exclusão de Manaus também pode limitar as inferências.

Concluindo, recomenda-se cautela ao analisar dados de gestantes (ou de quaisquer indivíduos) quando as características de interesse não podem ser observadas desde a sua origem, pois diferenças no tempo de observação podem ser cruciais. Ainda, a escolha da análise estatística mais apropriada deve ser baseada não apenas nos dados, mas também no tipo de delineamento. Uma vez que as técnicas computacionais, como a regressão de Cox, já estão bem desenvolvidas e amplamente disponíveis, até mesmo em pacotes estatísticos gratuitos, e não deveria haver empecilho para a realização das análises mais adequadas para cada tipo de estudo, evitando resultados viesados.

Referências Bibliográficas

- (1) Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic Measures of the Course and Outcome Pregnancy. *Epidemiologic Reviews* 2002; 24(2):91-101.
- (2) Koren G, Pastuszak A, Shinya I. Drugs and Pregnancy. *New England Journal of Medicine* April 1998; 338(16):1128-1137.
- (3) Dal Pizzol TS. Risco e benefícios para o feto e recém-nascido de medicamentos utilizados na gestação: misoprostol e antianêmicos. Tese de Doutorado pela Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Epidemiologia. Porto Alegre, BR-RS, 2006.
- (4) Costa SH, Vessey. Misoprostol and illegal abortion in Rio de Janeiro, Brazil. *Lancet* 1993; 341:1258-1261.
- (5) Coêlho HLL, Teixeira AC, Santos AP, Forte EB, Morais SM, La Vecchia C et al. Misoprostol and illegal abortion in Fortaleza, Brazil. *Lancet* 1993; 341:1261:1263.
- (6) Goldberg AB, Greenber BS, Darney PD. Misoprostol and pregnancy. *New England Journal of Medicine* 2001; 344:38-46.
- (7) Mengue SS, Dal Pizzol TS. Misoprostol, aborto e malformações congênitas. *Brás Ginecol Obstet.* 2008; 30(6):271-273.
- (8) Gallo MF, Cahill S, Castleman L, Mitchell MH. A systematic review of more than one dose of misoprostol after mifepristone for abortion up to 10 weeks of gestation. *Contraception* 2006; 74:36-41.
- (9) Marions L. Mifepristone dose in the regimen with misoprostol for medical abortion. *Contraception* 2006; 74:21-25.
- (10) Howards PP, Hertz-Picciotto I, Poole C. Conditions for Bias from Differential Left Truncation. *American Journal of Epidemiology* 2007; 165(4):444-452.
- (11) Horovitz DDG, Júnior JCL, Mattos RA. Atenção aos defeitos congênitos no Brasil: panorama atual. *Cadernos de Saúde Pública* 2005; 21(4):1055-1064.
- (12) Cecatti JG, Aquino MMA. Causas e Fatores Associados ao Óbito Fetal. *Revista de Ciências Médicas* 1998; 7(2):43-48.
- (13) Cnattingius S, Stephansson O. The Epidemiology of stillbirth. *Seminars in perinatology* February 2002; 26(1):25-30.

- (14) Cnattingius S, Haglund B, Kramer MS. Differences in late fetal death rates in association with determinants of small for gestational age fetuses: population based cohort study. *British Medical Journal* 1998; 316:1483-1487.
- (15) Aguiar MJB, Campos AS, Aguiar RALP, Lana ANA, Magalhães RL, Babeto LT. Defeitos de fechamento do tubo neural e fatores associados em recém-nascidos vivos e natimortos. *Jornal de Pediatria* 2003; 79(2):129-134.
- (16) Castro MLS, Cunha CJ, Moreira PB, Fernández RR, Garcias GL, Martino-Roth MG. Frequência das malformações múltiplas em recém-nascidos na Cidade de Pelotas, Rio Grande do Sul, Brasil, e fatores sócio-demográficos associados. *Cadernos de Saúde Pública* 2006; 22(5): 1009-1015.
- (17) Basso O, Olsen J, Christensen K. Recurrence Risk of Congenital Anomalies – The Impact of Paternal, Social and Environmental Factors: A population-based Study in Denmark. *American Journal of Epidemiology* 1999; 150:598-604.
- (18) Schüller L, Pastuszak AL, Sanseverino MTV, Orioli IM et al. Pregnancy outcome after exposure to misoprostol in Brazil: a prospective, controlled study. *Reproductive Toxicology* 1999; 13(2):147-151.
- (19) Pastuszak AL, Schüller L, Speck-Martins CE, Coelho KEFA, Cordello SM, Vargas F et al. Use of misoprostol during pregnancy and Möbius' Syndrome in infants. *New England Journal of Medicine* 1998; 338:1881-1885.
- (20) Parada CMGL, Pelá NTR. Idade Materna como Fator de Risco: Estudo com Primigestas na Faixa Etária Igual ou Superior a 28 anos. *Revista Latino-americana de enfermagem* 1999; 7(4):57-64.

Tabela 1 - Frequências absolutas (%) das características sócio-demográficas e clínicas das 4251 gestantes do EBDG incluídas na análise.

Característica	Sem anomalia	Com anomalia	Total
Centro			
Porto Alegre	1066(97,5)	27(2,5)	1093 (25,7)
Salvador	820 (97,8)	18 (2,2)	838 (19,7)
Fortaleza	1016 (98,8)	12 (1,2)	1028 (24,2)
Rio de Janeiro	457 (97,2)	13 (2,8)	470 (11,1)
São Paulo	807 (98,2)	15 (1,8)	822 (19,3)
Idade da mãe			
20 – 34 anos	3597 (98,1)	70 (1,9)	3667 (86,3)
35 – 48 anos	569 (97,4)	15 (2,6)	584 (13,7)
Cor de pele			
Não branca	2259 (97,7)	52 (2,3)	2311 (54,4)
Branca	1907 (98,3)	33 (1,7)	1940 (45,6)
Escolaridade da mãe			
Menos de 8 anos	1905 (97,7)	44 (2,3)	1949 (45,8)
8 a 11 anos	885 (98,4)	14 (1,6)	899 (21,2)
12 anos ou mais	1376 (98,1)	27 (1,9)	1403 (33,0)
Altura da mãe			
Até 1,55m	2139 (98)	43 (2)	2182 (51,5)
1,55 – 1,78m	2013 (98)	42 (2)	2055 (48,5)
Número de Filhos			
Nenhum	1125 (97,9)	24 (2,1)	1149 (30,5)
Mais de um	2575 (98,1)	50 (1,9)	2625 (69,5)
IMC pré-gravídico			
Menor que 19,8	632 (97,7)	15 (2,3)	647 (15,9)
19,8 a 26,0	2526 (98,4)	42 (1,6)	2568 (63,0)
26,1 a 29,0	476 (97,7)	11 (2,3)	487 (11,9)
Maior que 29,0	367 (97,3)	10 (2,7)	377 (9,2)
Idade gestacional da primeira consulta			
0 a 91 dias	1225 (98,0)	25 (2,0)	1250 (30,8)
92 a 182 dias	2272 (98,1)	43 (1,9)	2315 (57,0)
Mais de 182 dias	481 (97)	15 (3,0)	496 (12,2)
Distúrbios Hipertensivos			
Não	3700 (98,1)	70 (1,9)	3770 (88,7)
Eclâmpsia	107 (95,5)	5 (4,5)	112 (2,6)
HAS	359 (97,3)	10 (2,7)	369 (8,7)
Número de consultas de pré-natal			
0 a 5	1966 (97,9)	43 (2,1)	2009 (49,5)
6 a 14	2012 (98,1)	40 (1,9)	2052 (50,5)
Idade gestacional do recém-nascido			
37 semanas ou mais	3518 (98,2)	64 (1,8)	3582 (84,4)
< de 37 semanas	639 (96,8)	21 (3,2)	660 (15,6)
Peso do recém-nascido			
2500g ou mais	3708 (98,2)	67 (1,8)	3775 (90,7)
< 2500g	373 (95,9)	16 (4,1)	389 (9,3)
Hipertensão prévia			
Não	3833 (98,1)	76 (1,9)	3909 (91,9)
Sim	333 (97,4)	9 (2,6)	342 (8,1)
Aborto prévio			
Não	2678 (98,1)	53 (1,9)	2731 (72,4)
Sim	1022 (98)	21 (2)	1043 (27,6)
Uso de misoprostol			
Não	4070 (98,1)	80 (1,9)	4150 (97,6)
Sim	96 (95,1)	5 (4,9)	101 (2,4)

Tabela 2 – Associação entre anomalia congênita com o uso de misoprostol estimada pelos modelos univariáveis e multivariáveis.

Característica	Regressão logística RC (IC 95%)		Cox sem Truncamento HR (IC 95%)		Cox com Truncamento HR (IC 95%)	
	Univariável	Multivariável	Univariável	Multivariável	Univariável	Multivariável
Centro						
Porto Alegre	1	1	1	1	1	
Salvador	0,9 (0,47-1,59)	0,9 (0,50-1,68)	1,3 (0,68-2,28)	1,0 (0,53-1,95)	1,4 (0,68-2,92)	
Fortaleza	0,5 (0,24-0,93)	0,5 (0,23-0,96)	0,4 (0,19-0,74)	0,3 (0,13-0,55)	0,4 (0,20-0,91)	
Rio de Janeiro	1,1 (0,57-2,20)	1,2 (0,58-2,28)	8,2 (3,78-17,92)	12,1 (5,38-27,20)	Não estimável	
São Paulo	0,7 (0,39-1,39)	0,8 (0,43-1,56)	0,7 (0,37-1,35)	0,8 (0,40-1,49)	0,76 (0,36-1,61)	
Idade da mãe						
20 – 34 anos	1		1		1	1
35 – 48 anos	1,4 (0,77-2,38)		1,6 (0,91-2,80)		2,3 (1,19-4,36)	2,3 (1,15-4,46)
Cor de pele						
Não branca	1		1	1	1	
Branca	0,8 (0,48-1,17)		0,7 (0,45-1,09)	0,6 (0,37-1,00)	0,6 (0,31-0,98)	
IMC pré-gravídico						
Menor que 19,8	1		1			
19,8 a 26,0	0,7 (0,39-1,27)		0,6 (0,34-1,13)		0,6 (0,27-1,16)	
26,1 a 29,0	1,0 (0,44-2,14)		1,0 (0,45-2,11)		1,0 (0,41-2,61)	
Maior que 29,0	1,2 (0,51-2,58)		1,2 (0,55-2,75)		0,9 (0,32-2,74)	
Distúrbios Hipertensivos						
Não	1		1		1	
Eclâmpsia	2,5 (0,98-6,24)		2,4 (0,98-6,02)		3,0 (1,07-8,33)	
HAS	1,5 (0,75-2,88)		2,3 (1,16-4,38)		1,6 (0,58-4,51)	
Número de consultas de pré-natal						
6 a 14	1		1	1	1	1
0 a 5	1,1 (0,71-1,70)		2,4 (1,54-3,81)	2,9 (1,79-4,58)	2,4 (1,37-4,27)	2,3 (1,26-4,05)
Idade gestacional do recém-nascido						
37 semanas ou mais	1		1	1	1	1
< de 37 semanas	1,8 (1,10-2,98)		3,1 (1,86-5,18)	2,5 (1,37-4,49)	3,3 (1,67-6,46)	2,3 (1,04-5,02)
Peso do recém-nascido						
2500g ou mais	1	1	1	1	1	1
< 2500g	2,4 (1,36-4,14)	2,3 (1,29-3,96)	3,8 (2,18-6,77)	2,2 (1,14-4,07)	3,9 (1,83-8,43)	2,5 (1,08-5,98)
Uso de misoprostol						
Não	1	1	1	1	1	1
Sim	2,7 (1,05-6,69)	2,9 (1,12-7,44)	3,0 (1,19-7,45)	3,3 (1,28-8,43)	4,2 (1,49-12,01)	4,8 (1,69-13,79)

7. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo a comparação de três diferentes tipos de análise estatística: regressão logística, regressão de Cox sem truncamento à esquerda e regressão de Cox com truncamento à esquerda. É conhecido que a regressão logística está consagrada na literatura, tanto pela sua abrangência em diversas situações, como pela facilidade e simplicidade de modelar e entender as análises desejadas. Porém, a análise de sobrevivência ainda tem sido pouco explorada, principalmente no que diz respeito a estudos com gestantes.

Particularmente, a técnica de regressão de Cox com truncamento à esquerda ainda foi pouco explorada na literatura se comparada com o modelo que não considera o truncamento. Parece interessante difundir esse tipo de modelagem uma vez que muitos estudos referentes à gestação são publicados anualmente e esta técnica pode revelar novos achados ou a comprovação de hipóteses antigas. Neste sentido, é importante incentivar a utilização de ferramentas de análise que consideram o truncamento à esquerda.

Embora já tenha sido ressaltado por alguns autores, mesmo que de forma mais teórica, é crucial evidenciar o fato de que as gestantes são arroladas em épocas diferentes, e o tempo, no que diz respeito à gravidez é fundamental. Diversos eventos durante a gestação podem ser dependentes da idade gestacional e sua não consideração pode produzir estimativas viesadas, associações espúrias ou, eventualmente, a não identificação de associações com preditores importantes. Uma característica do modelo logístico é a incapacidade de acomodar essas diferenças no tempo e, conseqüentemente, não deveria ser utilizada nestas situações devido à possibilidade de vieses. Porém, a maioria dos estudos sobre gestantes encontrados na literatura utilizou regressão logística e, talvez, se os dados fossem analisados de outra forma poderiam evidenciar outros achados. Ainda, no modelo de regressão de Cox sem truncamento as gestantes são vistas como arroladas no mesmo ponto do tempo, fato que também pode produzir resultados viesados, embora potencialmente menores do que com a regressão logística. Isto ocorre porque embora as gestantes sejam tratadas de maneira igual no ponto inicial, como no modelo logístico, o desfecho considera que o tempo de acompanhamento pode ser diferente entre as gestantes. Por outro lado, o modelo de regressão de Cox com truncamento à esquerda consegue lidar com essas entradas em tempos diferenciados e, claro, fazer o acompanhamento ao longo do tempo. Assim, em resumo, este modelo trata as gestantes de maneira diferenciada não só no início do estudo, bem como ao longo deste.

Os modelos ajustados para o desfecho anomalia congênita foram diferentes tanto quanto aos preditores quanto à magnitude das estimativas da associação com o uso referido de misoprostol. De um modo geral, os resultados das análises de regressão de Cox com o truncamento à esquerda parecem

mais coerentes. Primeiro porque as mulheres não foram acompanhadas desde o momento da concepção e muitas delas realizaram tardiamente sua primeira consulta de pré-natal para iniciar os cuidados com a gravidez. A principal influência que o arrolamento de grávidas em clínicas de pré-natal pode trazer é a perda de informação do que acontece no período entre a concepção e o arrolamento, e não é correto considerar que todas tenham sido arroladas no mesmo momento da gestação. Segundo, é importante observar que a utilização do misoprostol possivelmente possibilitou que muitas mulheres abortassem e, sendo assim, nem ao menos tiveram a possibilidade de serem arroladas.

Considerando as evidências que apontam para a teratogenicidade do misoprostol, se a gestante fez uso do medicamento e não conseguiu abortar o feto, este acabou sujeito ao risco de desenvolver alguma anomalia congênita. Sendo assim, a estimativa de efeito significativa encontrada entre a ingestão do medicamento e o desfecho de anomalia congênita no modelo com truncamento parece ser a mais acurada se comparada com as demais análises. Embora o intervalo de confiança encontrado possa ser considerado amplo, isto pode ser visto como consequência do pequeno número de casos encontrados na amostra. As anomalias se desenvolvem em períodos diferenciados da gestação e, portanto, tratando as gestantes de maneira diferenciada ao longo do tempo pode gerar estimativas mais acuradas.

É importante assinalar que os resultados do modelo de regressão de Cox com truncamento à esquerda estão em concordância com a maioria dos relatos encontrados na literatura. Como são diferentes dos resultados do modelo de Cox sem truncamento à esquerda e de regressão logística, evidencia que a escolha do modelo implica sim em conclusões diferenciadas, particularmente quanto à magnitude das associações.

Concluindo, recomenda-se muito cuidado ao analisar dados de gestantes (ou de quaisquer indivíduos) quando as características de interesse não podem ser observadas desde a sua origem, pois diferenças no tempo de observação podem ser cruciais, uma vez que o desfecho e a exposição podem ser diferenciados ao longo do tempo. Os pesquisadores interessados em analisar esse tipo de dado deveriam considerar outros tipos de modelagem, além da regressão logística, como forma de incrementar suas análises e buscar resultados mais fidedignos. A escolha da análise estatística mais apropriada deve sempre ser baseada não só nos dados, mas também no tipo de delineamento, porém, uma vez que algumas variáveis possam ser recuperadas deve-se sempre preferir aqueles modelos que dêem conta das características e particularidades dos dados.

Uma vez que as técnicas e os recursos computacionais, como a regressão de Cox, já estão bem desenvolvidas e amplamente disponíveis, até mesmo em pacotes estatísticos gratuitos, esta questão não deveria ser empecilho para a realização das análises mais adequadas para cada tipo de estudo.

Resumindo, os resultados encontrados ao longo desse estudo dão conta que:

- existe associação entre o uso de misoprostol e anomalia congênita;

- a idade da mãe, o número de consultas de pré-natal, a idade gestacional e o peso do recém nascido são, de forma geral, preditores importantes para anomalia congênita. Assim, mulheres que procuram um atendimento mais seqüencial podem reduzir o risco desfechos não desejados da gestação, como anomalia congênita;

- os valores de razão de chances se apresentaram, em geral, inferiores aos valores de razão de azares, o que pode ser indício de subestimação dos parâmetros na regressão logística; e,

- a amplitude dos intervalos de confiança pode ser explicada pelo pequeno número de gestantes que manifestaram o desfecho.

8 ANEXO: PROJETO DE PESQUISA

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
Programa de Pós-Graduação em Epidemiologia**

**Risco de morte fetal e anomalia congênita por misoprostol
através do modelo de Cox com truncamento à esquerda**

Autor: Marília Canabarro Zordan

Professor Orientador: Prof. Dr. Álvaro Vigo

Co-orientadora: Tatiane da Silva Dal Pizzol

Colaboradores: Sotero Serate Mengue e Maria Inês Schmidt

1 Justificativa

O uso de medicamentos durante a gravidez exige precauções quanto à escolha do fármaco e dosagem apropriada devido às mudanças fisiológicas da unidade materno-feto-placentária acarretando variações nos processos de absorção, distribuição e eliminação de medicamentos (1). Entretanto, o principal foco da atenção em relação ao uso de medicamentos na gestação diz respeito à interferência na formação, crescimento e desenvolvimento do embrião e do feto, com destaque para os efeitos teratogênicos (2).

Nem sempre os estudos pré-comercialização podem fornecer informações dos possíveis danos que um medicamento pode causar, como por exemplo, a sua teratogenicidade, uma vez que existem dificuldades metodológicas para sua realização, como a extrapolação dos resultados em animais para seres humanos ou então devido ao princípio ético da não-maleficência em gestantes, crianças e outros indivíduos. Sendo assim, os estudos epidemiológicos pós-comercialização são, normalmente, imprescindíveis nessa área. Estudos de coorte e caso-controle, principalmente, têm sido uma alternativa importante não apenas na avaliação do risco teratogênico, mas também no estabelecimento da efetividade do medicamento utilizado em condições clínicas usuais, em populações específicas (2,3).

O uso de medicamentos na gravidez pode ser agrupado em: profilático, terapêutico e não-terapêutico. Entre os usos profiláticos, o mais difundido é a suplementação de sais minerais e vitaminas. Compõe o uso terapêutico o tratamento de distúrbios agudos ou crônicos, relacionados ou não à gestação. O uso não-terapêutico é definido aqui como o emprego de medicamentos com indicações diversas daquelas estabelecidas pelo fabricante ou com finalidades abusivas, tais como o uso de medicamentos com efeito estimulante ou depressor do sistema nervoso central e de medicamentos para diagnóstico ou interrupção da gravidez (3,4).

O uso de medicamentos e outros produtos com ação farmacológica para induzir a menstruação têm sido associado ao diagnóstico de gravidez ou à regularização do ciclo menstrual. O principal motivo relacionado a essa prática é, provavelmente, provocar o aborto, principalmente em países onde é proibido, como é o caso do Brasil (3,5). Em particular, o misoprostol, um fármaco licenciado para o tratamento de úlcera péptica, tornou-se rapidamente popular pelo seu potencial abortivo, mencionado na bula do produto. Estudos confirmaram a suspeita do largo uso do medicamento entre as mulheres brasileiras. Pesquisas realizadas em todo país revelaram percentuais de uso de misoprostol entre 50 e 75%, isolado ou associado a outros medicamentos (6,7,8). O risco à saúde materna não é o único problema associado ao uso indiscriminado de substâncias abortivas. Ensaio clínico controlados

demonstram que a falha abortiva com o uso de misoprostol em abortos médicos pode ultrapassar 10%, exigindo o emprego de método abortivo adicional para consumação do procedimento (9). Sob condições controladas proporcionadas por ensaios clínicos, o uso de misoprostol apresenta taxas de falhas abortivas indesejáveis. Nas situações em que o aborto ilegal é comumente realizado, sem assistência médica adequada e com emprego de doses inadequadas, o misoprostol pode não interromper a gestação, gerando ansiedade quanto ao risco para o feto. Alguns estudos demonstram uma associação entre o uso de misoprostol com morte fetal, aborto espontâneo e anomalia congênita (3).

Muitos estudos sobre aborto utilizam a regressão logística como uma maneira de analisar a relação entre o desfecho e os fatores de exposição, além de estimar as medidas de associação. Porém, este método não leva em consideração que as mulheres entram no estudo em tempos diferentes, uma vez que, em geral, as mulheres são arroladas a partir de clínicas de pré-natal, ou seja, após a concepção. Estes tipos de eventos são truncados a esquerda. Ainda, algumas gestações podem terminar antes mesmo da mulher ser recrutada.

Truncamento à esquerda surge quando a amostragem da população é incompleta, como ocorre com os sujeitos que não conseguem ficar tempo suficiente sem manifestar o evento de interesse para serem observados, de tal forma que o investigador pode nem estar ciente da existência desses indivíduos (10).

Muitos estudos sobre aborto espontâneo são truncados a esquerda porque as mulheres são recrutadas de clínicas de pré-natal após a concepção (que seria o tempo de origem ideal para um estudo de acompanhamento de mulheres grávidas). Sendo assim, as mulheres entram no estudo em diferentes idades gestacionais e, mais importante, uma parte da população fonte é desconhecida por causa dos abortos espontâneos ocorridos antes do arrolamento. Esta situação dá as mulheres que levaram a gravidez até o nascimento uma maior oportunidade de entrar no estudo. Calcular os efeitos estimados sem levar em conta o truncamento a esquerda pode introduzir vieses porque a entrada no estudo pode ser diferente tanto quanto ao status do desfecho ou da exposição (11). O modelo de regressão logística usado para a avaliação do risco de perda da gravidez não permite incorporar truncamento à esquerda. Em contraste, o modelo de Cox considera o tempo sob observação como desfecho, permitindo, assim, incorporar o truncamento à esquerda através de uma série de conjuntos de risco que incluem somente as mulheres que estão sob observação no tempo em que ocorre o evento (11).

Estes motivos revelam a importância de estimar o risco da exposição ao misoprostol durante a gestação através de um método que permita considerar ou atenuar o impacto da parte da população desconhecida devido aos abortos ocorridos antes do arrolamento.

2 Objetivos

Geral

- Estimar o risco de morte fetal ou mal-formação para a exposição ao misoprostol durante a gestação através do modelo de Cox com truncamento à esquerda no Estudo Brasileiro do Diabetes Gestacional (EBDG).

Específico

- Comparar os riscos estimados para a morte fetal por exposição ao misoprostol através da regressão logística, modelo de Cox tradicional e o modelo de Cox com truncamento à esquerda.

3 Métodos

Vários artigos mencionam o truncamento à esquerda, mas não examinam especialmente como isto altera os riscos. Na análise com o modelo de Cox que acomoda o truncamento à esquerda, as gestantes contribuem para a verossimilhança condicional à entrada em observação naquela determinada idade gestacional, em oposição à regressão de Cox tradicional, onde todas gestantes contribuem para a verossimilhança no início do estudo. Esta modificação da verossimilhança pode afetar a relação entre a razão de azares (hazard ratio) e a razão de chances (11).

Na construção da função de verossimilhança para dados truncados, precisa-se considerar com cuidado que informações cada observação nos fornece. Uma observação correspondendo a um tempo exato do evento fornece informações da probabilidade de que o evento ocorra naquele tempo que é aproximadamente igual à função de densidade de X naquele tempo. Para dados truncados à esquerda, estas probabilidades são substituídas pelas probabilidades condicionais apropriadas. Mais especificamente, a verossimilhança para estes dados devem incorporar o componente $f(x)/S(Y)$, onde $f(x)$ é a função densidade de probabilidade (12).

Para derivar os estimadores da função de sobrevivência para dados truncados à esquerda e censurados à direita, é preciso lembrar que para calcular as estatísticas em tempos anteriores é necessário o número de indivíduos que entraram no estudo antes do tempo t_i e que tem um tempo de estudo de pelo menos t_i , ou seja, define-se um Y_i como o número de indivíduos com uma idade maior que t_i e com um tempo de falha menor que ele. Sendo assim, pode-se usar o estimador padrão para a

função de sobrevivência, proposta por Kaplan-Meier, chamada de estimador Produto-Limite ou o estimador baseado no estimador Produto-Limite de Nelson-Aalen do risco acumulado (12).

Porém, deve-se tomar cuidado ao interpretar essas estatísticas. Por exemplo, o estimador Produto-Limite da função de sobrevivência no tempo t é agora um estimador da probabilidade de sobrevivência além de t , condicional a sobrevivência para o menor tempo de entrada. Alguns cuidados ao aplicar diretamente estes estimadores devem ser tomados, pois dependendo os valores de tempo de sobrevivência e sobreviventes o estimador Produto-Limite pode ser nulo para todos os t_i além deste ponto, mesmo que sejam observadas sobrevivências e falhas além deste ponto. Nestes casos, é comum estimar a função de sobrevivência condicional a sobrevivência de certo tempo onde o estimador Produto-Limite não pode ser igual a zero devido à consideração somente daqueles tempos de falhas além deste ponto (12).

Os dados foram obtidos do Estudo Brasileiro sobre Diabetes Gestacional (EBDG), uma coorte composta por 5564 mulheres com 20 anos ou mais, que se encontravam entre as semanas 21 e 28 de gestação e que se apresentaram para a consulta em centros de atendimento pré-natal do SUS, localizados em seis capitais brasileiras (Porto Alegre, São Paulo, Rio de Janeiro, Salvador, Fortaleza e Manaus). As gestantes foram arroladas de forma consecutiva de acordo com a procura por atendimento nos centros no período de fevereiro de 1991 a junho de 1995. Detalhes do estudo foram previamente publicados por Mengue et al (5,13).

Os desfechos são aborto espontâneo ou morte fetal e má formação, a principal variável preditora é o uso do misoprostol, ao passo que centro, peso ao nascer, nascimento pré-termo e fumo são considerados possíveis potenciais confundidores. O modelo de Cox com truncamento a esquerda será ajustado através do procedimento PHREG do programa SAS e através do programa R (14,15).

4 Aspectos Éticos

O EBDG obteve aprovação dos comitês de ética dos respectivos centros onde o estudo foi realizado. Neste estudo, será assinado termo de compromisso para a utilização de dados secundários provenientes do banco de dados do EBDG.

5 Cronograma

TAREFAS	2007	2008		2009	
		1° Semestre	2° Semestre	Jan/Fev	Março
Revisão da literatura					
Desenvolvimento de rotinas computacionais					
Teste das rotinas					
Defesa do projeto					
Redação da dissertação e artigo					
Defesa preliminar					
Sessão pública					

6 Referências

- (1) Loebstein R, Koren G. Clinical relevance of therapeutic drug monitoring during pregnancy. *Therapeutic Drug Monitoring*. 2002; 24(1):15-22
- (2) Koren G, Pastuszak A, Shinya I. Drugs and Pregnancy. *New England Journal of Medicine* April 1998; 338(16):1128-1137.
- (3) Dal Pizol TS. Risco e benefícios para o feto e recém-nascido de medicamentos utilizados na gestação: misoprostol e antianêmicos. Tese de Doutorado pela Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Epidemiologia. Porto Alegre, BR-RS, 2006.
- (4) Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Ações Programáticas Estratégicas. Pré-natal e puerpério. Atenção qualificada e humanizada. Manual técnico. Brasília: Ministério da Saúde, 2005: 67-68.
- (5) Mengue SS, Schenkel, Duncan BB, Schmidt MI. Prevalence and Clinical Correlates of Unsuccessful Use of Drugs to Induce Menstruation. *Contraception* 1998; 57:93-97.
- (6) Coêlho HLL, Teixeira AC, Santos AP, Forte EB, Morais SM, La Vecchia C et al. Misoprostol and illegal abortion in Fortaleza, Brazil. *Lancet* 1993; 341:1261:1263.

- (7) Fonseca W, Misago C, Correia LL, Parente JAM, Oliveira FC. Determinantes do aborto provocado entre mulheres admitidas em hospitais em localidades da região Nordeste do Brasil. *Revista de Saúde Pública* 1996; 30(1):279-286.
- (8) Fonseca W, Misago C, Freitas P, Santos E, Fernandes L, Correia L. Características sócio-demográficas reprodutivas e médicas de mulheres admitidas por aborto em hospital da Região Sul. *Cadernos de Saúde Pública* 1998; 14(2):279-286
- (9) Goldberg AB, Greenber BS, Darney PD. Misoprostol and pregnancy. *New England Journal of Medicine* 2001; 344:38-46.
- (10) Guang G. Event-History Analysis for Left-Truncated Data. *Sociological Methodology*, Peter V. Marsden; 1993:217-243.
- (11) Howards PP, Hertz-Picciotto I, Poole C. Conditions for Bias from Differential Left Truncation. *American Journal of Epidemiology* 2007; 165(4):444-452.
- (12) Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. New York, Springer-Verlag 1996.
- (13) Mengue SS, Schenkel EP, Duncan BB, Schmidt MI. Uso de medicamentos por gestantes em seis cidades brasileiras. *Revista de Saúde Pública* 2001; 35(5):415-420.
- (14) Klein JP, Pelz CJ. *Analysis of Survival Data: A Comparison of Three Major Statistical Packages (SAS, SPSS, BMDP)*. 1996.
- (15) Lai HJ, Cheng Y, Cho H, Kosorok MR, Farrel PM. Association between Initial Disease Presentation, Lung Disease Outcomes, and Survival in Patients with Cystic Fibrosis.