

# **COMPARAÇÃO DE RESULTADOS DE UMA COORTE SOB AS ABORDAGENS PROSPECTIVA E HISTÓRICA: AMAMENTAÇÃO NO PRIMEIRO ANO DE VIDA**

GIZELTON PEREIRA ALENCAR

Dissertação apresentada à Faculdade de  
Saúde Pública da Universidade de São  
Paulo para obtenção do grau de Mestre.

Área de concentração: Epidemiologia

ORIENTADOR: PROF. DR. JOSÉ MARIA  
PACHECO DE SOUZA

São Paulo

2003

## **Agradecimentos**

Os meus sinceros agradecimentos ao meu orientador José Maria Pacheco de Souza que, com paciência e atenção, auxiliou-me nos meus passos.

À minha querida Célia, aos meus pais e irmãs: o apoio da família sempre esteve presente.

Às professoras Marcia Furquim e Denise Bergamaschi que sempre deram força na vida e no meio acadêmico.

Ao Departamento que ofereceu estrutura para estudar e aos funcionários que ajudavam sempre que possível.

À banca examinadora pelas valiosas contribuições.

À FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pela bolsa de estudo concedida (processo 00/04081-5).

## Resumo

Alencar GP. **Comparação de resultados de uma coorte sob as abordagens prospectiva e histórica: amamentação no primeiro ano de vida.** São Paulo; 2003. [Dissertação de Mestrado - Faculdade de Saúde Pública da USP].

**Objetivo.** Uma coorte de crianças foi observada sob duas abordagens: coorte prospectiva e coorte retrospectiva (histórica) e o objetivo foi comparar as estimativas da função de riscos do modelo de Cox entre as duas abordagens e a mesma comparação com o modelo complemento log-log. Além disso, comparar as estimativas da função de riscos pelos modelos de Cox e complemento log-log para cada uma das abordagens, separadamente. **Métodos.** Quando se estuda o tempo de sobrevivência da amamentação sem as covariáveis, foram obtidas as estimativas pela técnica atuarial e modelo complemento log-log da informação retrospectiva e a técnica de Kaplan-Meier para a informação prospectiva. Os modelos de Cox e complemento log-log foram utilizados para estimar a razão de riscos (HR) com covariáveis para as duas abordagens. **Resultados.** Sem as covariáveis, a comparação entre as duas abordagens mostrou que as estimativas de  $S(t)$  pela informação retrospectiva estão um pouco defasadas em relação à medida prospectiva. Com as covariáveis, os resultados dos modelos de Cox e complemento log-log são semelhantes tanto para os dados prospectivos quanto para os retrospectivos. Foram semelhantes, também, os resultados de um mesmo modelo para cada uma das fontes de informação. **Conclusões.** Em geral, as estimativas foram bastante próximas em quaisquer das comparações. Somente a variável hábito de fumar da mãe durante a gravidez permaneceu nos modelos finais para todas as técnicas utilizadas, com estimativas próximas, reforçando semelhança entre as várias abordagens.

Descritores: estudos longitudinais, Kaplan-Meier, modelo de Cox, modelo complemento log-log, amamentação.

## Summary

Alencar GP. **Comparação de resultados de uma coorte sob as abordagens prospectiva e histórica: amamentação no primeiro ano de vida.** [Comparison between results of a cohort under prospective and historical approaches: breastfeeding in the first year of age]. São Paulo (BR); 2003. [Dissertation - Faculdade de Saúde Pública da Universidade de São Paulo, Brazil].

**Objective.** A cohort of children was observed under two approaches: prospective cohort and retrospective (historical) cohort and the objective was to compare the estimates of the hazard ratio from Cox model between the two approaches and do the same comparison using the complementary log-log model. Moreover, to compare the estimates of the hazard ratio from the two models to each one of the approaches, separately. **Methods.** When the survival time for breastfeeding is studied without the covariates, the comparison between the results of actuarial technique and the complementary log-log model has been made for the recorded information. The Kaplan-Meier technique has been used with the daily notebook measures. The Cox and complementary log-log models can estimate the risk rate of covariates categories to both approaches. **Results.** Without the covariates, the comparison between the two information resources showed that retrospective measures give lower estimates than that from the prospective measures. With covariates, the estimates are not so different and led to the same results. **Conclusions.** The estimates of each one of the comparisons were too close. Just the variable mother's smoking during the pregnancy stayed in the final models for every techniques used, with close estimates, reinforcing likeness between the several approaches.

Descriptors: longitudinal studies, Kaplan-Meier, Cox model, complementary log-log model, breastfeeding.

# ÍNDICE

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>6</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>7</b>
3.1	A PESQUISA “ALIMENTAÇÃO NO PRIMEIRO ANO DE VIDA”	7
3.2	VARIÁVEIS DE ESTUDO	10
3.3	MÉTODOS DE ANÁLISE DE SOBREVIDA	11
3.3.1	<i>A função de sobrevida e a função de risco</i>	11
3.3.2	<i>Modelos para dados agrupados</i>	12
3.3.2.1	Técnica atuarial ou da tábua de vida	12
3.3.2.2	Modelo complemento log-log	12
3.3.3	<i>Modelos para dados contínuos</i>	14
3.3.3.1	Modelo de Kaplan-Meier	15
3.3.3.2	Modelo de riscos proporcionais de Cox	15
3.4	PROCEDIMENTO DE OBTENÇÃO E VERIFICAÇÃO DOS MODELOS	17
3.5	QUESTÕES ÉTICAS	18
<b>4</b>	<b>RESULTADOS</b>	<b>18</b>
4.1	SEM USAR AS COVARIÁVEIS	18
4.1.1	<i>Comentários</i>	22
4.2	USANDO CADA COVARIÁVEL	23
4.2.1	<i>Caderno diário</i>	24
4.2.2	<i>Recordatório</i>	26
4.3	USANDO AS COVARIÁVEIS EM UM MODELO MULTIVARIADO	29
4.3.1	<i>Comentários</i>	39
<b>5</b>	<b>DISCUSSÃO</b>	<b>41</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>44</b>
<b>7</b>	<b>REFERÊNCIAS</b>	<b>45</b>
	<b>ANEXOS</b>	<b>49</b>
	ANEXO 1 – TERMO DE CONSENTIMENTO	50
	ANEXO 2 – FORMULÁRIO 1	52
	ANEXO 3 – FORMULÁRIO 2	56
	ANEXO 4 – FORMULÁRIO 3	58

ANEXO 5 – QUESTIONÁRIO RECORDATÓRIO .....	63
ANEXO 6 – CADERNO DIÁRIO .....	65
ANEXO 7 – ESCOLHA DOS MODELOS.....	67
ANEXO 8 – PROGRAMA NO STATA PARA O MODELO COMPLEMENTO LOG-LOG E ALTERAÇÕES NO BANCO DE DADOS.....	87
ANEXO 9 – PARECER DO COMITÊ DE ÉTICA EM PESQUISA DA FSP - COEP .....	90

# 1 Introdução

A alimentação no primeiro ano de vida de uma criança é importante para seu desenvolvimento satisfatório. Essa alimentação envolve a amamentação e a introdução de outros alimentos (RADRIGAN 1989; HITCHCOCK 1985).

Uma possível forma de caracterizar a amamentação das crianças desde o nascimento é por meio de um estudo epidemiológico do tipo coorte. Pode-se considerar duas abordagens. Uma, denominada coorte prospectiva, acompanha as crianças no decorrer do tempo e registra o instante em que cessa a amamentação, no momento de sua ocorrência. Outra abordagem depende da obtenção de informações já registradas ou de inquéritos que as tornem disponíveis, baseando-se na memória das pessoas (KLEINBAUM 1982). Pode-se chamar esta abordagem de coorte retrospectiva ou histórica.

## **Estudos prospectivos, retrospectivos e de corte transversal**

Os estudos epidemiológicos observacionais podem ter duas referências de tempo: transversal e longitudinal. O estudo do tipo coorte, tanto prospectivo como retrospectivo, é um estudo longitudinal.

Na coorte prospectiva, as medidas da exposição e demais covariáveis são tomadas antes da ocorrência do evento. No caso retrospectivo, as medidas da exposição, as demais covariáveis e a ocorrência do evento são avaliadas após a ocorrência do mesmo (PEREIRA 1999; ROTHMAN e GREENLAND 1998).

Embora sejam bastante usados, os termos prospectivo e retrospectivo podem ser encontrados com outras denominações. Na literatura, a coorte retrospectiva é também chamada de coorte histórica ou coorte prospectiva não concorrente. A coorte prospectiva é, por vezes, chamada de coorte prospectiva concorrente (KELSEY 1986; LILIENFELD 1976).

Em diversas situações, estudos de corte transversal podem fornecer informações retrospectivas a respeito da ocorrência de um evento e podem ser utilizadas como se fosse uma coorte retrospectiva. Em um estudo de coorte, existe a aleatorização dos indivíduos a participar da pesquisa, o que não ocorre, em muitos casos, com o estudo

de corte transversal. Os casos desse último estudo podem representar um viés de seleção e interferir nas estimativas produzidas no que se refere às relações causais que podem ser feitas em estudos de coorte (KLEINBAUM 1982).

### **Vieses**

Um possível viés que pode surgir em estudos que envolvem a anotação diária de um formulário é provocado pelo indivíduo participar de uma pesquisa. A pesquisa pode fazer com que o participante aumente a atenção para um determinado evento que se quer estudar, provocando uma alteração dos resultados. Assim, o fato de se fazer parte de uma pesquisa sobre alimentação, pode influenciar na conduta da amamentação da criança. Além disso, o entrevistador pode, também, causar um viés. A presença de um entrevistador pode, por exemplo, influenciar sistematicamente os resultados sobre a anotação como um todo. Já o viés de recordação, ou de memória, pode ocorrer mais em um estudo de abordagem retrospectiva (PEREIRA 1999). Uma comparação entre os resultados das abordagens prospectiva e histórica pode apontar diferenças devidas a vieses. Para que se considere a comparação como um instrumento para medir o viés de memória, deve-se supor que a abordagem prospectiva é a que mais se aproxima da medida real – dela é obtida a medida padrão. Essa suposição surge da idéia de que essa abordagem apresenta menos problemas de memória, pois o registro é feito no momento da ocorrência do evento, não se considerando os possíveis problemas no preenchimento desse registro.

### **Estudos de sobrevivência em coortes**

O tempo pode ser a variável de maior interesse em um estudo. De uma forma geral, nesses estudos existe o interesse de se identificar fatores que influenciam o tempo até a ocorrência do evento.

Em um estudo de coorte, o início de observação do indivíduo é feito a partir de um momento definido como, por exemplo, o nascimento. A partir daí, o indivíduo é observado até ocorrer um evento de interesse e quando esse acontece diz-se que ocorreu uma **falha**. Quando não é possível se observar a ocorrência de um evento, diz-se que ocorre uma **censura** na observação e é considerado o tempo até aquele



momento. Isso ocorre por algumas razões: o indivíduo ultrapassa o tempo limite estabelecido para o estudo sem que a falha ocorresse; ocorre algo que se impossibilita a continuação da observação (como o óbito, o abandono do estudo ou a perda de localização).

A **sobrevida** é definida como o tempo desde o início da observação do indivíduo no estudo até a ocorrência do evento de interesse (falha) ou até a censura. Define-se, a partir daí, a probabilidade de sobrevida acumulada a partir de um tempo  $t$  por meio da função probabilidade de sobrevida acumulada que para facilitar chamaremos de função de sobrevida. Uma outra função utilizada é a função de riscos que representa a taxa de falhas em um certo intervalo curto de tempo, dado que o indivíduo não teve falha em um momento anterior a esse intervalo. A partir dessa função, pode-se comparar duas funções de riscos a partir da medida da razão das funções de riscos de falha (HR: *hazard ratio*).

Em estudos epidemiológicos, a estimativa da função de sobrevida, da função de riscos e da razão de riscos das covariáveis estudadas é bastante utilizada e pode ser obtida por meio de técnicas estatísticas que serão apresentadas na Seção Material e Métodos com mais detalhes.

Duas técnicas são bastante utilizadas para se obter as estimativas da função de sobrevida: a técnica atuarial e a técnica de Kaplan-Meier. Por meio dessas técnicas, é possível estimar a sobrevida para um conjunto de indivíduos e considerar essas estimativas segundo categorias de uma covariável (KLEIN e MOESCHBERGER 1997; LEE 1992; LAWLESS 1982; KAPLAN e MEIER 1958).

Outras duas técnicas de análise de sobrevida são os modelos de Cox e complemento log-log, que podem ser utilizados para se avaliar o efeito de covariáveis em um modelo multivariado (HOSMER e LEMESHOW 1999; COLLETT 1994).

O modelo de Cox pressupõe riscos proporcionais e modela seus coeficientes por meio da verossimilhança parcial (COX 1972). O modelo complemento log-log estima os seus coeficientes maximizando-se a verossimilhança da forma usual (LAWLESS 1982) e considera, da mesma forma que no modelo de Cox, a proporcionalidade entre as funções de riscos de falha.

### **Análise de sobrevivência em estudos que envolvem o tempo de amamentação**

A amamentação é entendida, de acordo com a Organização Mundial da Saúde (OMS), como o processo pelo qual o lactente recebe leite materno independentemente de consumir outros alimentos (OPAS/OMS 1991).

Em estudos de amamentação, a técnica atuarial surge com bastante frequência para se calcular o tempo mediano de amamentação em estudos de coorte.

Rea et al. (1997) utilizaram a técnica atuarial para avaliar o efeito do trabalho formal, em uma coorte de mulheres, na duração da amamentação, da amamentação exclusiva e da predominante e observadas, respectivamente, as medianas de 150, 10 e 70 dias.

Kummer et al. (2000) utilizaram a técnica atuarial para obter a função de sobrevivência entre duas coortes (uma de 1987 e outra de 1994), comparando-as com o teste de *log-rank*, sem encontrar diferenças. Além disso, utilizaram a regressão logística para avaliar o efeito de covariáveis como renda, escolaridade, cor da mãe, assistência pré-natal e estado civil, ressaltando-se maior chance de deixar de amamentar mães de baixa renda e cor branca. O estudo utilizou curvas de sobrevivência atuarial considerando-se cortes, em meses, até 6 meses.

Assis et al. (1994) avaliaram o tempo de amamentação e de amamentação exclusiva de uma coorte provinda de comunidades rurais do semi-árido baiano e utilizou a técnica atuarial para obter estimativas da sobrevivência da amamentação até 1 ano de idade. A duração mediana da amamentação ficou em 90,28 dias e nenhuma das mães amamentou exclusivamente no peito durante mais do que um dia.

Hardy et al. (1982) avaliaram a amamentação em Paulínia, SP, e utilizaram a técnica atuarial para o tempo de amamentação e amamentação exclusiva obtendo 3,3 e 2,2 meses de tempo mediano, respectivamente. Utilizaram, ainda, a técnica atuarial separadamente por idade da mãe (até 27 anos; 28 e mais), sexo e lugar de residência (urbano; rural).

Passos et al. (2000) usaram a técnica atuarial para obter a duração mediana do aleitamento materno em Ouro Preto em um estudo de corte transversal.

Diversos artigos trabalham com a abordagem prospectiva e apresentam fatores de risco para o desmame. Alguns estudos têm o ponto de partida em hospitais (Scott

JA et al. 1999; Escamilla et al. 1997) e, no seguimento, as casas dos entrevistados foram visitadas até um determinado tempo, que varia de 1 a 12 meses. Nesses trabalhos, é utilizadas a técnica de Kaplan-Meier para se selecionar variáveis para posterior aplicação do modelo de Cox e se estudar, a partir dessa modelagem, os fatores associados à duração da amamentação.

Abada et al. (2001) estudaram por meio de um modelo de Cox o efeito de variáveis socioeconômicas, demográficas de saúde e alimentos suplementares. A modelagem foi feita por meio de um modelo hierarquizado levando-se em consideração uma estrutura de causalidade a partir de fatores sócio-econômicos, de alimentos suplementares e variáveis relacionadas à saúde. As variáveis mais importantes foram nível de educação da mãe, número de filhos tidos, ocupação da mãe, cuidado pré-natal e idade de introdução de fórmulas infantis. Neste trabalho, a suposição da proporcionalidade dos riscos foi verificada para a variável de localidade (urbano; rural) e como essa suposição não foi satisfeita, o modelo utilizado foi então estratificado.

Rutishauser et al. (1992) detectaram por meio do modelo de Cox o efeito das seguintes covariáveis: hábito de fumar da mãe na gravidez, idade materna, IMC, um escore relacionado às atividades da mãe e momento do primeiro aleitamento.

Levy et al. (2000) avalia principalmente o efeito de chupeta, chupar o dedo, além da atenção e cuidado dispensados à criança sobre a duração da amamentação.

No estudo de Gigante et al. (2000), a análise estimou a prevalência de amamentação aos 6 meses e a associou às covariáveis por meio do modelo de Cox e de regressão logística. Idade materna até 20 anos, não ter tido filhos anteriormente, hábito de tabagismo na gestação e idade gestacional até 37 semanas estiveram relacionadas à maior chance de desmame.

Weiderpass et al. (1998) avaliaram uma coorte durante três meses e usou o modelo de Cox para avaliar o tipo de parto (normal, cesariana eletiva, cesariana emergencial), a classe social e a educação materna, mas não apresentaram significância estatística.

No levantamento bibliográfico para o presente trabalho, não foi encontrado nenhum trabalho que utilize o modelo complemento log-log em amamentação. Serão

citados alguns exemplos, aplicados a outros tipos de dados, para ilustrar o uso dessa técnica.

Chalita et al. (2000) exemplificaram o uso do modelo complemento log-log para estimar a função de sobrevivência quanto à resistência de mangueira períodos de seca, acompanhadas em diversos anos entre 1973 a 1992.

Lancelot et al. (2002) utilizaram o modelo complemento log-log para estimar o efeito de algumas variáveis na mortalidade de fêmeas de cabras em cinco coortes diferentes, acompanhadas até 1 ano de vida.

Kosterman et al. (2000) acompanharam uma amostra de uma coorte de 808 jovens, com entrevistas anuais nas idades de 10 a 16 e 18 anos, para identificar as covariáveis relacionadas ao tempo de início do uso de álcool e maconha. Foi utilizado o modelo complemento log-log e foram consideradas covariáveis dependentes do tempo.

No presente estudo foram utilizados dados da pesquisa “Alimentação no primeiro ano de vida” (FAPESP 96/06073-2) que gerou um conjunto de dados com informações de um caderno de registro diário e de um recordatório. O caderno diário gera as informações da coorte prospectiva e o recordatório atua como fonte de informações do passado para simular a coorte retrospectiva.

Supõe-se que estudos de coorte prospectiva tenham boa confiabilidade quanto à informação obtida, pois a observação estaria mais próxima da ocorrência do evento. Por outro lado, pesquisas com coortes retrospectivas são mais rápidas e mais baratas, mas não se conhece o quanto seus resultados se aproximam dos obtidos em um estudo com coorte contemporânea de fato. A proposta deste trabalho é fazer um estudo que compare os resultados dessas duas abordagens, a partir de informações obtidas de um mesmo grupo de mães/crianças.

## **2 Objetivos**

Comparar os resultados da abordagem de coorte retrospectiva com os resultados de coorte prospectiva, usando as técnicas de análise de sobrevivência atuarial e de Kaplan-Meier e os modelos de risco proporcional de Cox e complemento log-log,

com as informações de amamentação no primeiro ano de vida. Mais especificamente, objetiva-se:

1) Obter as estimativas da função de sobrevivência, aplicando as técnicas atuarial e de Kaplan-Meier, considerando-se fonte de informação retrospectiva e fonte prospectiva, respectivamente;

2) comparar as estimativas da função de risco, obtidas pelo modelo proporcional de Cox, ao se utilizar uma fonte de informação retrospectiva e uma fonte prospectiva;

3) comparar as estimativas da função de risco, obtidas pelo modelo complemento log-log, ao se utilizar uma fonte de informação retrospectiva e uma fonte prospectiva;

4) comparar os resultados obtidos com o uso dos modelos de riscos proporcionais de Cox e complemento log-log, com a fonte de informação retrospectiva;

5) comparar os resultados obtidos com o uso dos modelos de riscos proporcionais de Cox e complemento log-log, com a fonte de informação prospectiva.

### **3 Material e métodos**

#### **3.1 A pesquisa “Alimentação no primeiro ano de vida”**

Na Faculdade de Saúde Pública da Universidade de São Paulo foi desenvolvida a pesquisa “Alimentação no primeiro ano de vida”. É um estudo de coorte com recém-nascidos selecionados na maternidade do Hospital Universitário (HU) da Universidade de São Paulo, situado na Cidade Universitária no Município de São Paulo, seguidos desde o nascimento até completarem o primeiro ano de vida. A coleta de dados da coorte foi realizada para crianças nascidas no período de outubro de 1998 a outubro de 1999. Como o acompanhamento de cada dupla mãe/criança teve duração de um ano, a coleta de informações foi concluída em outubro de 2000.

No HU nascem, em média, sete crianças por dia, cujas mães podem ser classificadas como da comunidade USP (alunas, funcionárias docentes e não docentes, dependentes de funcionários) e de fora da comunidade (residentes na região do bairro do Butantã).

Estudantes de nutrição da Faculdade de Saúde Pública e outras pessoas indicadas pela participação anterior em outras pesquisas foram treinadas para fazer o contato e dar orientações às mães em todas as etapas da pesquisa.

A seleção das mães era feita um dia após o nascimento dos bebês conforme registrado na ficha de nascimentos, enquanto as mães ainda estavam na enfermaria. Eram entrevistadas as primeiras quatro mães que satisfizessem ao protocolo, que consistia em nascimentos de crianças sem intercorrências, mães residentes no município de São Paulo e que concordassem em participar da pesquisa assinando um termo de consentimento (Anexo 1). Não foram incluídas no estudo crianças com malformação congênita, nascidas de parto gemelar, que necessitassem de internação hospitalar por um período superior a cinco dias, com morte materna ou com a mãe necessitando internação por mais de cinco dias.

O Formulário 1 (Anexo 2), preenchido na entrevista inicial no HU, quando a mãe/criança era selecionada, apresenta características da criança, da gestação, do parto e da família. O Formulário 2 (Anexo 3) era preenchido pela mãe na primeira visita, aos 15 dias de idade da criança, e trata de condições do domicílio e posse de bens para a caracterização sócio-econômica.

No Formulário 3 (Anexo 4) eram anotadas informações sobre o crescimento e cuidados da saúde (onde se incluem situações de imunização), consultas médicas, razões para o desmame quando este ocorresse, além de medidas antropométricas e de desenvolvimento da criança. Esse formulário era aplicado em todas as visitas.

No Anexo 5 encontra-se uma ficha, utilizada pela entrevistadora nas visitas de 12 meses, que contém perguntas relativas à recordação da mãe (recordatório) no que se refere à amamentação e outras medidas como peso e comprimento da criança ao nascer e aos 6 meses.

Para o seguimento, as mães foram entrevistadas no HU e receberam um conjunto inicial de fichas formatadas para leitura ótica com informações diárias (Anexo 6) no qual elas próprias, dia a dia, marcavam os eventos relacionados à alimentação da criança. Esse conjunto de fichas é chamado de caderno diário. Foram, também, programadas visitas domiciliares de entrevistadores para as idades de 15, 30, 60, 90, 180, 270 e 360 dias da criança. Nessas visitas, o entrevistador recolhia o

caderno diário preenchido pela mãe, dava um novo caderno diário e obtinha informações adicionais sobre a alimentação, o crescimento, desenvolvimento e saúde da criança, em cada período, mediante o preenchimento de questionários estruturados. Cada ficha ótica permitia o registro de 60 dias de eventos (Anexos de 2 a 4).

O Esquema 1 apresenta o fluxo da coorte. As perdas ocorridas durante o seguimento foram motivadas pela não aceitação da mãe em continuar no estudo (35), não-preenchimento do primeiro caderno diário mesmo com a presença dos cadernos seguintes (19), falecimento da criança até 15 dias (1), internação da criança até 15 dias (1).

Esquema 1. Dinâmica da coorte do início ao término da pesquisa



Para ser possível a comparação entre os resultados foram consideradas 360 pares de mães/crianças que possuem as informações completas até 360 dias, nas duas fontes de informação, para a variável que se refere ao tempo de amamentação. Não foram consideradas, portanto, as situações em que alguma das fontes de informação não estava preenchida.

### 3.2 Variáveis de estudo

As variáveis consideradas e suas categorias foram:

- sexo da criança;
- tipo de parto:  
normal, fórceps, cesárea;
- cor da pele:  
branca, não branca (preta, parda/mulata, amarela/oriental);
- idade da mãe:  
até 19 anos (jovem), de 20 a 34 anos (ideal); acima de 34 anos (madura);
- escolaridade da mãe:  
universitário, secundário, fundamental;
- condição marital da mãe:  
casada, não casada;
- hábito de fumar da mãe:  
sim, não;
- o pai da criança morar com a mãe:  
sim, não;
- posse de bens:  
indicador de posse de nenhum (0 itens, categorizado como "não"), um ou mais itens (pelo menos 1 item, categorizado como "sim"): computador, carro, televisão paga e telefone;
- amamentação e tempo de amamentação:  
No caderno, exemplificado no Anexo 6, o tempo de amamentação foi obtido por meio da contagem dos alvéolos marcados pela mãe. Admitiu-se que todas as crianças ao nascer iniciaram a amamentação. O final da amamentação corresponde à última marcação após pelo menos três marcações (que podem ou não ter interrupções entre elas) em um período de 60 dias no campo "Leite de peito", independentemente da frequência anterior observada.

O questionário recordatório de 12 meses está apresentado no Anexo 5. O campo que corresponde à amamentação está representado na questão "Até quando recebeu leite de peito?" que pode ser preenchido com o tempo em meses ou em dias.



### 3.3 Métodos de análise de sobrevivência

A duração da amamentação é a principal variável resposta a ser considerada e gera a variável aleatória  $T > 0$ , que representa o tempo até a ocorrência do desmame. O tempo observado pode representar: o tempo de ocorrência do evento; o tempo em que o indivíduo esteve em observação sem acontecer o evento; ou o tempo de ocorrência do evento dado na forma de um intervalo.

Os dados provenientes do caderno podem apresentar um tipo de censura denominado censura à direita e indica que a informação sobre o tempo de amamentação só é conhecida até um determinado momento e não se sabe o que acontece com a amamentação além desse tempo.

Os dados do recordatório apresentam um tipo de censura denominado censura por intervalos e está representando o intervalo de tempo no qual o término da amamentação ocorreu. Pode-se, ainda, ter casos de censura à direita no recordatório representados por aqueles questionários que não apresentam resposta porque a amamentação ainda não se encerrou até o momento da pergunta da entrevista. A informação produzida pelo recordatório é chamada intervalar ou agrupada.

#### 3.3.1 A função de sobrevivência e a função de risco

A função probabilidade de sobrevivência acumulada (chamada aqui de função de sobrevivência),  $S(t)$ , é definida como a probabilidade de o tempo  $T$  ser maior que um dado tempo  $t$ , e está representada por

$$S(t) = P(T > t).$$

Uma outra função que pode ser utilizada é a função distribuição dos tempos de falha e é dada por

$$F(t) = P(T \leq t) = 1 - S(t),$$

que indica a probabilidade de o evento ocorrer até o tempo  $t$ .

A função de risco é definida por

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t},$$

e representa a taxa de falha instantânea no tempo  $t$ .

Essa função é também chamada de *hazard function*, *hazard rate*, taxa de falha ou força de mortalidade, para cada tempo  $t$ .

A probabilidade de ocorrer a falha no intervalo  $[t, t+\Delta t)$  pode ser, então, dada por  $h(t)\Delta(t)$  (COLLETT 1994).

### 3.3.2 Modelos para dados agrupados

Muitas vezes, não é possível precisar o tempo em que o evento de falha (fim da amamentação) ou a censura ocorreu, mas se conhece o respectivo intervalo de tempo de sua ocorrência. Esses dados são chamados de agrupados.

#### 3.3.2.1 Técnica atuarial ou da tábua de vida

A técnica *atuarial* (LEE 1992; LAWLESS 1982) ou da *tábua de vida*, de certa maneira limitada quanto a testes de hipóteses, estima  $S(t) = P(T > t)$ , mediante a equação

$$\hat{S}(t_i) = \prod_{j=1}^i (1 - \hat{\theta}_j^A), \text{ onde}$$

$i$  representa o número de intervalos considerados;

$\hat{\theta}_j^A$  é o estimador da probabilidade de ocorrer o evento no intervalo  $j$  e pode ser escrito como

$$\hat{\theta}_j^A = \frac{d_j}{n_j - \frac{1}{2}w_j}, \text{ onde}$$

$d_j$  representa o número de falhas no intervalo  $j$ ;

$n_j$  é o número de indivíduos em risco no início do intervalo  $j$ ;

$w_j$  é o número de indivíduos censurados no intervalo  $j$ .

Nessa representação, supõe-se que as censuras ocorrem uniformemente no intervalo. O denominador  $n_j - \frac{1}{2}w_j$  é chamado número efetivo de indivíduos expostos a sofrer falha.

#### 3.3.2.2 Modelo complemento log-log

O modelo complemento log-log é um modelo de regressão que trabalha com resposta binária. Utilizar este modelo viabiliza considerar dados de censura intervalar

em modelo de riscos proporcionais (PRENTICE e GLOECKER 1978). Esse modelo é também denominado modelo de riscos proporcionais com dados agrupados (HOSMER e LEMESHOW 1999; LAWLESS 1982).

Consideramos  $n$  indivíduos com os tempos  $T_i$ ,  $i = 1, \dots, n$ , agrupados em  $k$  intervalos  $I_j = [a_{j-1}, a_j)$ ,  $j = 1, \dots, k$  e seja  $\mathbf{x}_i$  o vetor de covariáveis do indivíduo  $i$ . Seja  $D_j$  o conjunto de indivíduos que falharam no intervalo  $I_j$ ; e  $R_j$ , o conjunto de indivíduos que estavam em risco no intervalo  $I_j$ .

A probabilidade do  $i$ -ésimo indivíduo falhar até  $a_j$  dado que não falhou até  $a_{j-1}$  é

$$\theta_{ij} = \theta_j(\mathbf{x}_i) = P(T_i \leq a_j \mid T_i \geq a_{j-1}, \mathbf{x}_i).$$

Reescrevendo-se em termos da função de sobrevivência, temos

$$\theta_j(\mathbf{x}_i) = 1 - P(T_i \geq a_j \mid T_i \geq a_{j-1}, \mathbf{x}_i) = 1 - \frac{S(a_j \mid \mathbf{x}_i)}{S(a_{j-1} \mid \mathbf{x}_i)},$$

onde  $\frac{S(a_j \mid \mathbf{x}_i)}{S(a_{j-1} \mid \mathbf{x}_i)}$  representa a probabilidade do  $i$ -ésimo indivíduo não falhar até

$a_j$ , dado que não falhou até  $a_{j-1}$ .

A função de verossimilhança é dada por:

$$\ell(\beta) = \prod_{j=1}^k \prod_{i \in R_j} \left[ (\theta_j(\mathbf{x}_i))^{c_{ij}} (1 - \theta_j(\mathbf{x}_i))^{(1-c_{ij})} \right],$$

onde  $c_{ij}$  é variável indicadora da ocorrência de falha para o indivíduo  $i$  no intervalo  $j$ .

Essa é a verossimilhança para variáveis com distribuição Bernoulli com probabilidade de sucesso (no caso, de ocorrência do evento)  $\theta_j(\mathbf{x}_i)$  e variável resposta  $c_{ij}$ .

A idéia é permitir usar modelos de regressão binária para modelar uma regressão com riscos proporcionais para dados com censura por intervalo.

Mostra-se algebricamente que

$$\theta_j(\mathbf{x}_i) = 1 - \exp\{-\exp(\mathbf{x}_i\beta + \tau_j)\}, \text{ onde}$$

$$\tau_j = \ln \left[ -\ln \left( \frac{S_0(a_j)}{S_0(a_{j-1})} \right) \right] \text{ representa a razão entre } S_0 \text{ do início e } S_0 \text{ do fim do}$$

intervalo  $i$  e representa o efeito de intervalo de tempo.

Assim, o modelo de regressão é dado por:

$$\ln[-\ln(1 - \theta_j(\mathbf{x}_i))] = \mathbf{x}_i' \boldsymbol{\beta} + \tau_j$$

Esse modelo é denominado *complemento log-log* (CARSTENSEN 1996; FARRINGTON 1996). Neste modelo, dados com censura intervalar podem ser modelados considerando o problema como se fosse um problema de regressão com resposta binária. No contexto de modelos lineares generalizados (McCULLAGH e NELDER 1989), esse modelo corresponde a um modelo de regressão da família binomial com função de ligação complemento log-log.

Deve-se notar que na construção do banco de dados é necessário que, para cada indivíduo em cada intervalo, exista uma linha que indique o resultado para a resposta  $c_{ij}$ , definida anteriormente. Ela indica se o evento ocorreu ou não no intervalo  $j$ . Por exemplo, o indivíduo que tem falha no terceiro intervalo terá três linhas no banco de dados: as duas primeiras onde constam a resposta  $c_{ij} = 0$  e a terceira, a resposta  $c_{ij} = 1$ . O indivíduo que é censurado no intervalo 2 tem duas linhas no banco de dados, ambos com  $c_{ij} = 0$ . O programa no STATA que promove a construção do banco da forma descrita e um exemplo do resultado desse programa estão no Anexo 8.

Para cada intervalo  $j$ , o estimador de  $S_0$  em  $t_j$ , é

$$\hat{S}_0(t_j) = \hat{S}_0(t_{j-1}) \exp(-\exp(\hat{\tau}_j)).$$

Para o primeiro intervalo,  $I_1$ ,  $\hat{S}_0(t_1) = \exp(-\exp(\hat{\tau}_1))$  e para o último intervalo  $I_7$ ,  $\hat{S}_0(t_7) = \hat{S}_0(t_6) \exp(-\exp(\hat{\tau}_7))$ .

Para obter estimadores considerando as covariáveis, o estimador fica como no modelo de riscos proporcionais

$$\hat{S}(t_j, \mathbf{x}_i, \hat{\boldsymbol{\beta}}) = [\hat{S}_0(t_{j-1})]^{\exp(\mathbf{x}_i' \hat{\boldsymbol{\beta}})}.$$

### 3.3.3 Modelos para dados contínuos

O tempo pode ser tratado como variável contínua dependendo da frequência de sua observação como, por exemplo, quando se observa dia a dia.

Será utilizado um índice que representa a ordenação dos tempos de falha. Primeiramente, os tempos de falha de cada um dos  $n$  indivíduos são ordenados e tem-se o primeiro tempo de falha, o segundo tempo de falha, e assim por diante:

$$T_{(1)} < T_{(2)} < T_{(3)} < \dots < T_{(j-1)} < T_{(j)} < \dots < T_{(n)}, \text{ onde}$$

$T_{(j)}$  identifica o j-ésimo tempo de falha que ocorre no decorrer do período.

### 3.3.3.1 Modelo de Kaplan-Meier

*Kaplan e Meier* (KLEIN e MOESCHBERGER 1997; KAPLAN e MEIER 1958) estimam a função de sobrevida  $S(t_{(j)})$  por meio da equação

$$\hat{S}(t_{(j)}) = \hat{S}(t_{(j-1)}) \hat{P}[T > t_{(j)} | T \geq t_{(j)}],$$

onde a segunda parte da expressão representa uma estimativa da probabilidade de sobrevida até o tempo ordenado de falha anterior ( $t_{(j-1)}$ ) multiplicado pela probabilidade condicional de sobrevida até o tempo de falha ordenado  $t_{(j)}$ , dada a sobrevida até pelo menos o tempo  $t_{(j)}$ .

A função de sobrevida estimada para  $S(t)$ , que indica a probabilidade de o tempo de sobrevida ser superior a um determinado tempo  $t$ , pode ser escrita como

$$\hat{S}(t) = \prod_{t_{(j)} \leq t} \left( 1 - \frac{d_j}{n_j} \right) = \prod_{t_{(j)} \leq t} \frac{n_j - d_j}{n_j} = \prod_{t_{(j)} \leq t} (1 - \hat{\theta}_j),$$

em que  $n_j$  é o número de indivíduos em risco de falha no tempo  $t_{(j)}$ ;  $d_j$  é o número de falhas e  $\hat{\theta}_j$  é uma estimativa da probabilidade de ocorrer o evento no tempo  $j$ , dado que não ocorreu o evento até o tempo  $j - 1$ . No índice do somatório, utilizam-se os tempos ordenados que são menores do que  $t$ .

Deve-se notar que esse estimador considera os indivíduos com censura para o cálculo do número de indivíduos até o tempo  $j$ ,  $n_j$ .

Teste de hipóteses no qual se consideram as hipóteses de igualdade de duas curvas de sobrevida  $S_0$  e  $S_1$  ( $H_0: S_0(t) = S_1(t)$  x  $H_1: S_0(t) \neq S_1(t)$ , para todo o período) é feito mediante o teste *log-rank* (PETO e PETO 1972) ou, de forma mais geral, o teste de Wilcoxon generalizado (BRESLOW 1970; GEHAN 1965), que propõe pesos diferentes para diferentes instantes de falha e engloba o teste *log-rank*.

### 3.3.3.2 Modelo de riscos proporcionais de Cox

O modelo de riscos proporcionais de Cox (COX 1972), tem a forma

$$h(t, \mathbf{x}, \beta) = h_0(t) \cdot HR,$$

em que  $h(t, \mathbf{x}, \boldsymbol{\beta})$ : função de risco (*hazard function*) dependente das covariáveis;  
 $h_0(t)$ : função de risco dependente somente do tempo;

$$HR = \exp\left(\sum_{i=1}^p \beta_i \mathbf{X}_i\right): \text{relação entre as funções de risco e que é dependente das}$$

covariáveis, representadas por  $\mathbf{x}$ , constante no tempo. Pode ser chamada de razão de risco de falhas (*hazard ratio*).

O modelo gera uma expressão para a relação de funções de riscos, em qualquer tempo  $t$ , entre indivíduos com as suas covariáveis  $(x_1, x_2, \dots, x_p)$  representadas por um vetor  $\mathbf{x}$ .

A verossimilhança desse modelo não é construída de forma usual. Foi proposta por Cox (1972) e é baseada na probabilidade de um indivíduo, com covariáveis  $\mathbf{x}$ , falhar em um determinado tempo  $t_j$  dada a ocorrência de falha em  $t_j$ . Considerando-se os tempos ordenados de falha  $t_{(1)} < t_{(2)} < \dots < t_{(j)} < \dots < t_{(n)}$  e  $R(t_{(j)})$  o conjunto de indivíduos em risco de falha ou censura até o momento imediatamente anterior a  $t_{(j)}$ , tem-se

$$\begin{aligned} &P[\text{indivíduo com covariáveis } x_{(j)} \text{ falhar em } t_{(j)} \mid \text{houve falha em } t_{(j)}] = \\ &= \frac{\exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})}{\sum_{i \in R(t_{(j)})} \exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})}. \end{aligned}$$

Tomando-se o produto das probabilidades para o total de  $d$  falhas, tem-se

$$L(\boldsymbol{\beta}) = \prod_{j=1}^d \frac{\exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})}{\sum_{i \in R(t_{(j)})} \exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})}.$$

Quando ocorrem empates nos tempos de falha, algumas aproximações para a verossimilhança podem ser feitas.

A mais simples de ser programada é a aproximação proposta por Breslow (1974) e é dada por

$$L(\boldsymbol{\beta}) = \prod_{j=1}^d \frac{\exp(\mathbf{s}_{(j)} \boldsymbol{\beta})}{\left(\sum_{i \in R(t_{(j)})} \exp(\mathbf{x}'_{(j)} \boldsymbol{\beta})\right)^{\delta_j}}, \text{ onde}$$

$\mathbf{s}_{(j)}$  é o vetor das somas de cada uma das covariáveis daqueles indivíduos que falharam no tempo  $j$ .

Efron propõe uma aproximação que é razoavelmente fácil de se programar, mas que apresenta problemas quando a proporção de censuras com relação ao número de falhas é muito grande (THERNEAU 2000).

Um método exato foi proposto por Kalbfleisch e Prentice (1980), mas necessita de grande quantidade de cálculos. Uma outra forma de calcular é por meio da verossimilhança média que é semelhante à aproximação proposta por Efron (THERNEAU 2000).

### **3.4 Procedimento de obtenção e verificação dos modelos**

As covariáveis utilizadas nos modelos, todas categóricas por natureza ou categorizadas, foram definidas de forma que a razão das funções de riscos (HR) aparecesse com valor maior que 1 para facilitar a interpretação. Foi levado em consideração, para algumas variáveis, o que diz a literatura a respeito de qual a categoria de referência a ser utilizada.

Para obtenção dos modelos, os procedimentos utilizados foram os seguintes:

1. Foram feitas as curvas de Kaplan-Meier para cada variável e aplicado o teste de *log-rank*;
2. Foram aplicados os modelos univariados;
3. Para selecionar quais variáveis seriam incluídas em um único modelo, foram selecionadas aquelas que possuísem um nível descritivo de até 15% para o teste de Wald dos modelos do passo anterior. Após ter definido o modelo, pode-se avaliar a permanência de cada uma das variáveis do modelo por meio do teste de razão de verossimilhanças. O nível de significância utilizado para a permanência das variáveis foi de 5%.
4. O modelo com todas as variáveis também foi feito para se observar as estimativas segundo as técnicas e fontes de informação.

No modelo de Cox, foi sempre utilizada a aproximação pelo método de Breslow para a estimação do modelo na ocorrência de empates entre os tempos de falha por ser disponível em muitos *softwares*.

Para se verificar a proporcionalidade dos riscos, foi utilizado um gráfico no qual se colocam os valores da função de sobrevivência obtidos pelo modelo e por um modelo

empírico. Para as estimativas do modelo de Cox foi utilizado como modelo de comparação o de Kaplan-Meier. Para as estimativas do modelo complemento log-log foi utilizado o modelo atuarial. Quanto mais próximos os valores da medida de Kaplan-Meier daqueles obtidos pelo modelo de Cox, mais se satisfaz a suposição de proporcionalidade dos riscos. Da mesma forma, isso pode ser visualizado para os valores do modelo complemento log-log e da técnica atuarial (HESS 1995).

### **3.5 Questões éticas**

O desenvolvimento do estudo está seguindo os requisitos da Resolução 196/96 do Conselho Nacional de Saúde/Ministério da Saúde do Brasil (COMISSÃO NACIONAL DE ÉTICA EM PESQUISA 2000) e as normas do Comitê de Ética da Faculdade de Saúde Pública da Universidade de São Paulo, que regulamentam pesquisas envolvendo seres humanos. O projeto foi avaliado e aprovado em 4 de junho de 2001 (Anexo 9).

## **4 Resultados**

### **4.1 Sem usar as covariáveis**

A comparação é feita por meio da observação das estimativas,  $\hat{S}(t)$ , segundo as diversas técnicas e fontes de informação. As tabelas dessa seção apresentam as estimativas  $\hat{S}(t)$  e respectivos intervalos de confiança de 95%.

Os resultados da técnica atuarial para os dados do recordatório de 12 meses estão apresentados na Tabela 1 que contém: os intervalos de tempo considerados; o número de crianças em amamentação no início de cada intervalo, o número de crianças que deixaram o leite de peito (falha), o número de observações perdidas de observação (censura), a função de sobrevivência estimada ( $\hat{S}(t)$ ), para o instante final do intervalo, o erro padrão de  $\hat{S}(t)$ , e o intervalo de 95% de confiança para  $\hat{S}(t)$ .



Tabela 1 – Estimativas de  $S(t)$  pela técnica atuarial – recordatório

Intervalo de tempo	Crianças em amamentação no início do intervalo	Falhas	Censuras	$\hat{S}(t)$	Erro padrão de $\hat{S}(t)$	IC 95%	
0 – 15	360	20	0	0,9444	0,0121	0,9152	0,9638
16 – 30	340	15	0	0,9028	0,0162	0,8633	0,9277
31 – 60	325	36	0	0,8028	0,0218	0,7510	0,8365
61 – 90	289	41	0	0,6889	0,0253	0,6250	0,7242
91 – 180	248	67	0	0,5028	0,0271	0,4473	0,5532
181 – 270	181	45	0	0,3778	0,0263	0,3297	0,4325
271 – 360	136	9	127	0,3309	0,0274	0,2890	0,3961

Note-se que o número de censuras no último intervalo não se refere a observações incompletas, mas a observações de crianças com 360 dias que ainda não deixaram o leite de peito. Refere-se ao tipo de censura tipo I.

Pelo recordatório de 12 meses, a sobrevida no intervalo terminado em 360 dias é 34,23% e a duração mediana da amamentação é de 90 dias.

O modelo que pode ser utilizado para a modelagem dos dados agrupados, como o produzido pelo recordatório, é o modelo complemento log-log, apresentado na Tabela 2.

Tabela 2 – Resultado da regressão complemento log-log – recordatório

intervalos	coeficiente	erro padrão	Z	$P> z $	IC 95%	
0 – 15	-2,861929	0,223637	-12,80	0	-3,300250	-2,423608
16 – 30	-3,098420	0,258221	-12,00	0	-3,604523	-2,592317
31 – 60	-2,142181	0,166762	-12,85	0	-2,469029	-1,815333
61 – 90	-1,877331	0,156326	-12,01	0	-2,183724	-1,570937
91 – 180	-1,155399	0,122675	-9,42	0	-1,395838	-0,914961
181 – 270	-1,252316	0,149579	-8,37	0	-1,545485	-0,959146
271 – 360	-2,681392	0,333398	-8,04	0	-3,334841	-2,027943

As estimativas de cada intervalo e as correspondentes funções de sobrevida estimadas  $\hat{S}(t)$  são apresentadas na Tabela 3.

Tabela 3 – Estimativas de  $S(t)$  pelo modelo complemento log-log – recordatório

Intervalo de tempo	Coefficiente de cada intervalo	$\hat{S}(t)$	IC 95%	
0 – 15	-2,861929	0,9444	0,9152	0,9638
16 – 30	-3,098420	0,9028	0,8492	0,9379
31 – 60	-2,142181	0,8028	0,7216	0,8618
61 – 90	-1,877331	0,6889	0,5862	0,7700
91 – 180	-1,155399	0,5028	0,3927	0,6011
181 – 270	-1,252316	0,3778	0,2677	0,4857
271 – 360	-2,681392	0,3528	0,2347	0,4687

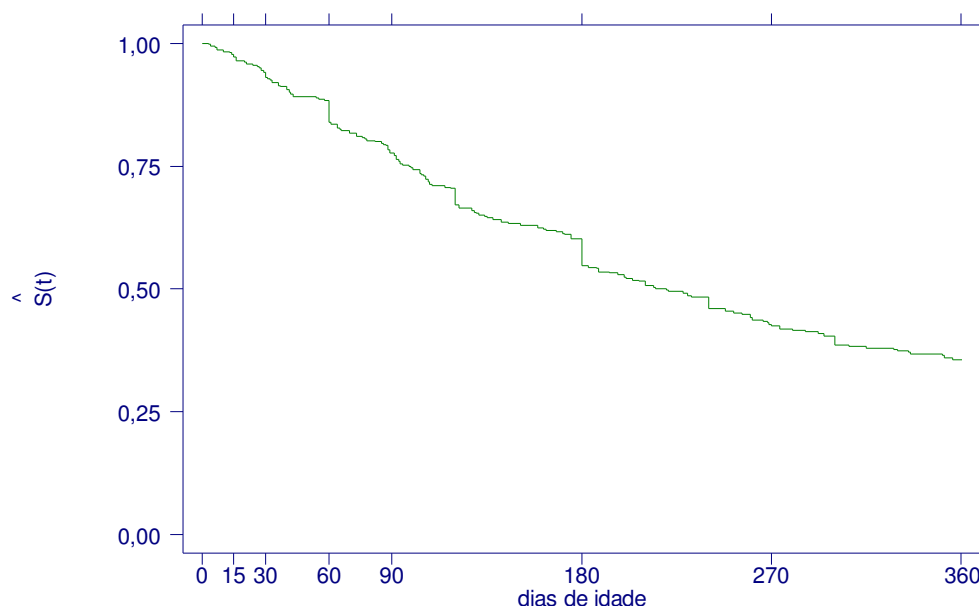
Como já era esperado, o modelo complemento log-log apresentou estimativas de  $S(t)$  idênticas à da técnica atuarial quando não existe censuras.

A técnica de Kaplan-Meier produz estimativas de  $S(t)$  em todo o momento que ocorre um evento de falha, o que é especialmente interessante no caso dos dados diários. Na Tabela 4, têm-se as estimativas somente para alguns tempos que são os pontos de corte dos dados recordatórios.

Tabela 4 – Estimativas de Kaplan-Meier – caderno diário

Tempo	Crianças em amamentação	Falhas acumuladas	Censuras acumuladas	$\hat{S}(t)$	Erro padrão de $\hat{S}(t)$	IC 95%	
0	360	-	-	-	-	-	-
15	360	20	0	0,9722	0,0087	0,9490	0,9850
30	340	15	0	0,9306	0,0134	0,8990	0,9525
60	325	36	0	0,8389	0,0194	0,7967	0,8730
90	289	41	0	0,7772	0,0220	0,7305	0,8168
180	248	67	0	0,5468	0,0265	0,4934	0,5970
270	181	45	0	0,4246	0,0264	0,3724	0,4758
360	136	9	127	0,3558	0,0258	0,3056	0,4063

Gráfico 1 – Gráfico de Kaplan-Meier – caderno diário



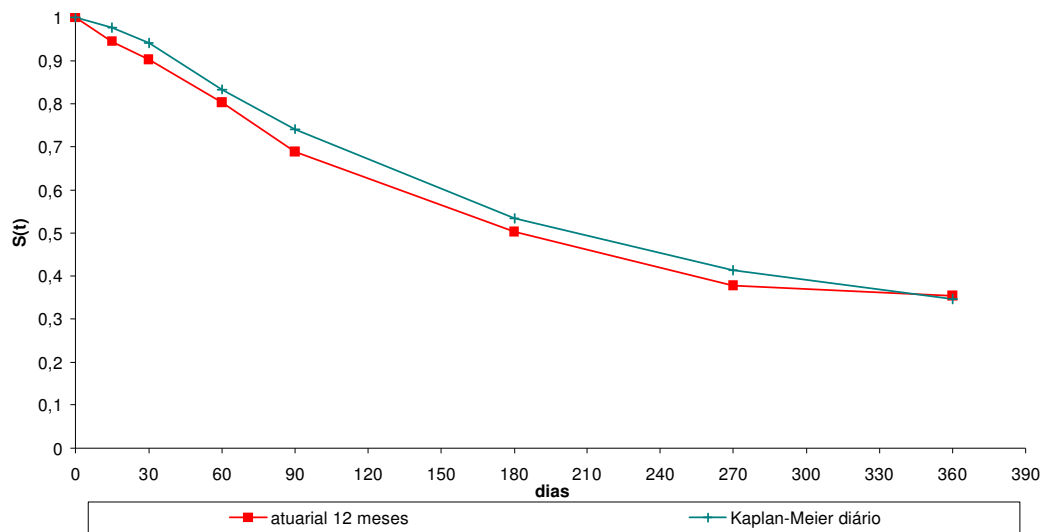
O Gráfico 1 apresenta a curva da proporção de crianças ainda em amamentação no decorrer do tempo, segundo a estimativa de Kaplan-Meier. Pode-se observar que existem algumas quedas mais pronunciadas nos tempos 60, 120, 180 e 240 dias. Note-se que o número de falhas corresponde ao ocorrido no tempo exato que está indicado em cada linha da Tabela 4.

A Tabela 5 e o Gráfico 2 apresentam o conjunto das estimativas de  $S(t)$ , segundo as três técnicas.

Tabela 5 – Estimativas de  $S(t)$ , segundo as técnicas e fontes de informação

Intervalo de tempo	recordatório		diário
	Técnica atuarial	Complemento log-log	Kaplan-Meier
0 – 15	0,9444	0,9444	0,9722
16 – 30	0,9028	0,9028	0,9306
31 – 60	0,8028	0,8028	0,8389
61 – 90	0,6889	0,6889	0,7772
91 – 180	0,5028	0,5028	0,5468
181 – 270	0,3778	0,3778	0,4246
271 – 360	0,3309	0,3528	0,3558

O Gráfico 2 mostra as estimativas para  $S(t)$  segundo as técnicas atuarial para o recordatório e técnica de Kaplan-Meier para o caderno diário.

Gráfico 2 – Estimativas de  $S(t)$ , segundo as técnicas e fontes de informação

Do início do período de observação até 270 dias, a função de sobrevivência estimada apresentada pelo recordatório é menor do que a calculada pelos dados do diário. Aos 360 dias, a estimativa para 12 meses, pela técnica atuarial também está bem próxima da estimativa obtida pelo diário.

#### 4.1.1 Comentários

O modelo de regressão com função de ligação complemento log-log é escrito na forma  $\ln[-\ln(1 - \theta_{ij})] = \mathbf{x}_i \boldsymbol{\beta} + \tau_j$  e a probabilidade condicional de falha é dada por  $\theta_{ij} = 1 - \exp\{-\exp(\mathbf{x}_i \boldsymbol{\beta} + \tau_j)\}$ .

Quando não são consideradas as covariáveis, ou seja, o vetor  $\mathbf{x}_i = \mathbf{0}$ , o modelo toma a forma simplificada  $\ln[-\ln(1 - \theta_{ij})] = \tau_j$  e  $\theta_{ij} = 1 - \exp\{-\exp(\tau_j)\}$ .

$$\text{Se } \tau_j = \ln \left[ -\ln \left( \frac{S_0(t_j)}{S_0(t_{j-1})} \right) \right], \quad \theta_{ij} = 1 - \exp\{-\exp(\tau_j)\} = 1 - \frac{S_0(t_j)}{S_0(t_{j-1})}.$$

$\theta_{ij}$  representa a probabilidade do indivíduo  $i$  falhar no intervalo  $j$ . Note-se que qualquer indivíduo, neste intervalo, tem a mesma probabilidade de falhar no intervalo  $j$ . Tem-se, então, a probabilidade com o índice referente ao intervalo  $j$ , independente de  $i$  para aquele intervalo,  $\theta_j$ , que pode ser interpretado como a proporção de falhas no intervalo  $j$ .

O estimador de  $\theta_j$  é semelhante ao da probabilidade de falhar no intervalo  $j$  na técnica atuarial baseada na proporção de falhas, representada por  $\hat{\theta}_j^A$ , e dada por

$$\hat{\theta}_j^A = \frac{d_j}{n_j - \frac{1}{2}w_j}.$$

Quando se estima  $\theta_j^A$ , o número de indivíduos que está em risco, no denominador, tem uma parte dos indivíduos censurados retirados, o que não ocorre no modelo complemento log-log que considera a censura no fim do intervalo. Assim, o estimador de  $\theta_j^A$  deve ser maior que o estimador de  $\theta_j$ . Portanto, as estimativas de  $S(t)$  serão menores pela técnica atuarial do que pelo modelo complemento log-log, na presença de censuras e, da mesma forma, serão iguais quando não houver censuras, que é o que ocorreu com os dados apresentados.

Quanto às respostas obtidas pelas fontes de informação, a comparação do tempo de sobrevivência entre os recordatórios e o caderno diário permite avaliar diferenças que poderiam expressar possível vício de memória. As estimativas de Kaplan-Meier para o caderno diário são, até 270 dias, sempre maiores do que as outras estimativas produzidas para os recordatórios.

Retornando ao Gráfico 1, a curva de Kaplan-Meier apresenta algumas quedas mais pronunciadas nos tempos 60, 120, 180 e 240 dias, que são múltiplos de 60. Isso é devido, possivelmente, ao fato de algumas mães não terem marcado corretamente o término da amamentação deixando para fazê-lo somente no começo da outra folha, pois cada uma das 6 folhas conta com 60 dias. Isso sugere que a curva, próximo aos dias mencionados, deveria ter um maior número de mães deixando de dar o leite de peito o que significa dizer que a curva mais próxima da realidade deve estar um pouco mais para baixo.

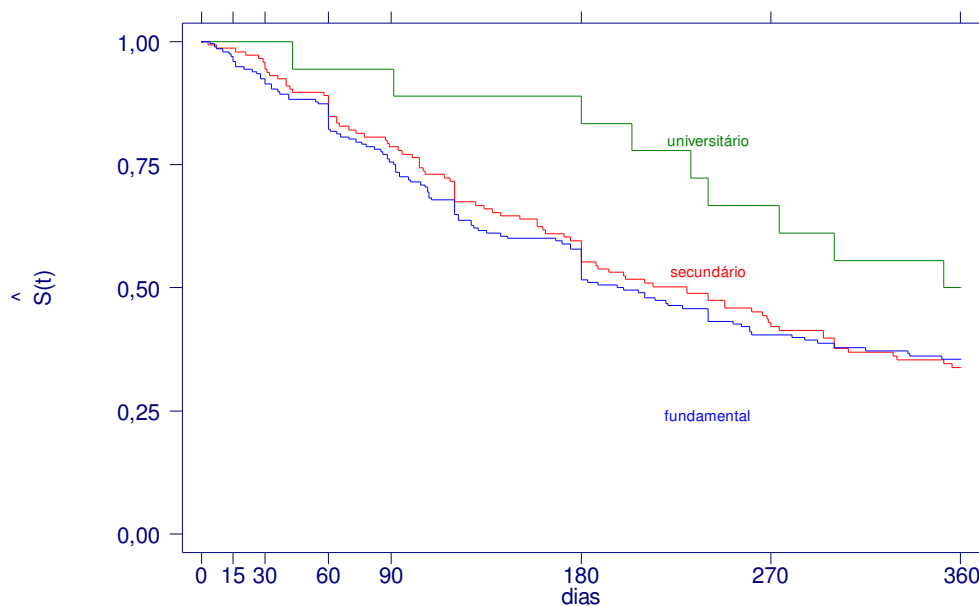
## 4.2 Usando cada covariável

A análise univariada pode ser utilizada para auxiliar na escolha de variáveis para inclusão em um modelo estatístico. Foram escolhidas algumas variáveis para apresentar os gráficos: idade da mãe, condição marital e hábito de fumar da mãe.

### 4.2.1 Caderno diário

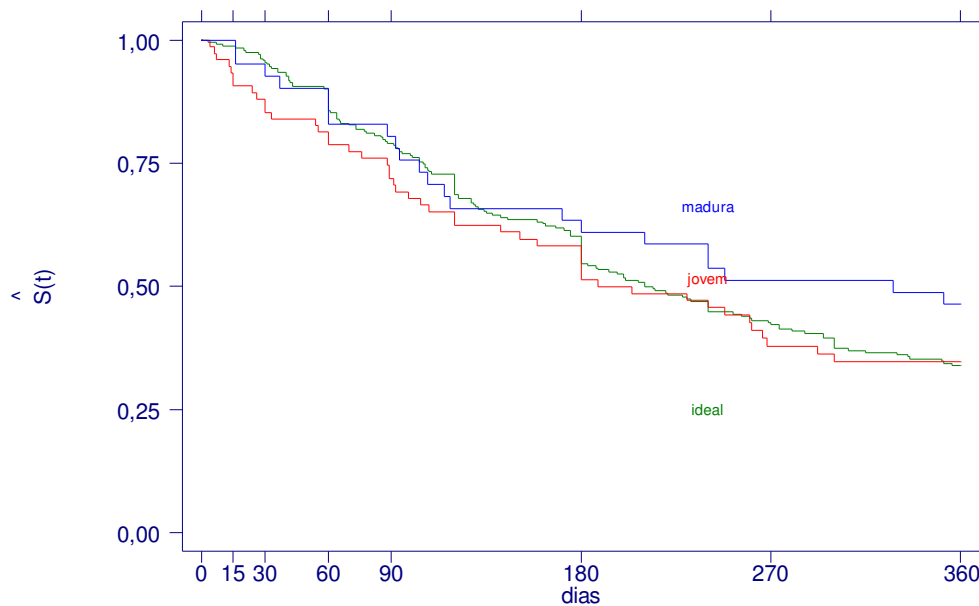
Os Gráficos 3 a 5 apresentam as curvas de Kaplan-Meier para as informações do caderno diário das variáveis escolhidas.

Gráfico 3 – Estimativas de Kaplan-Meier para a escolaridade materna – caderno diário



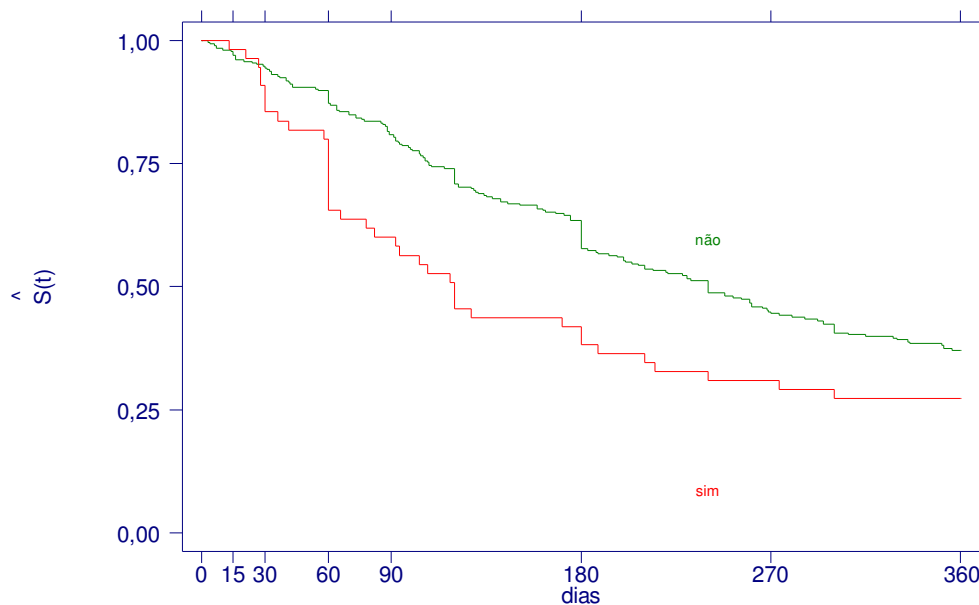
Para as curvas do Gráfico 3, o teste de *log-rank* indica não haver diferença entre as curvas de escolaridade materna ( $p=0,2218$ ). O número de universitários (18) é relativamente pequeno (5%).

Gráfico 4 – Estimativas de Kaplan-Meier para a idade materna – caderno diário



Para as curvas do Gráfico 4, o teste de *log-rank* indica não haver diferença entre as curvas de idade materna ( $p=0,3515$ ).

Gráfico 5 – Estimativas de Kaplan-Meier para hábito de fumar da mãe – caderno diário



Para as curvas do Gráfico 5, o teste de *log-rank* indica haver diferença entre as curvas de hábito de fumar ( $p=0,0156$ ).

Tabela 6 – Teste de *log-rank* – caderno diário

Variáveis	% de falhas	S(360)	<i>Log-rank</i>
<i>sexo</i>			
masculino	35,71	0,3364	0,7152
feminino	38,76	0,3751	
<i>tipo de parto</i>			
normal	35,00	0,3295	0,1867
fórceps	31,25	0,3051	
cesárea	45,83	0,4442	
<i>raça</i>			
branca	35,98	0,3447	0,9050
não branca	39,04	0,3738	
<i>idade da mãe (anos)</i>			
até 19 anos	35,66	0,3392	0,4133
de 20 ou 34	37,33	0,3471	
34 anos e mais	46,34	0,4634	
<i>escolaridade materna</i>			
nível universitário	50,00	0,5000	0,2566
nível secundário	35,86	0,3382	
nível fundamental	37,06	0,3552	
<i>condição marital</i>			
casada	37,46	0,3613	0,4487
não casada	36,36	0,3372	
<i>hábito de fumar</i>			
não	39,02	0,3704	0,0099
sim	27,27	0,2727	
<i>morar com o pai</i>			
sim	36,61	0,3514	0,9772
não	40,00	0,3801	
<i>posse de bens</i>			
sim	39,81	0,3790	0,1906
não	33,33	0,3208	

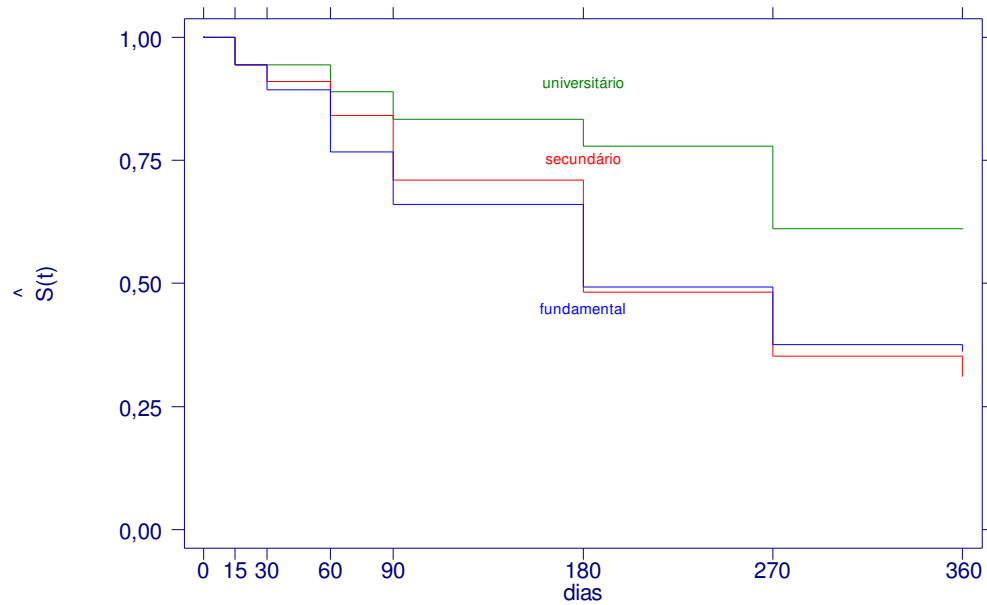
O Anexo 7 apresenta a análise univariada dos modelos de Cox e complemento log-log com cada uma das variáveis. Nelas, destacamos a presença das variáveis escolaridade da mãe e hábito de fumar da mãe durante a gestação, para os dois modelos.

#### 4.2.2 Recordatório

Os gráficos 6 a 8 apresentam as curvas de Kaplan-Meier para as informações do recordatório de 12 meses das variáveis escolhidas.

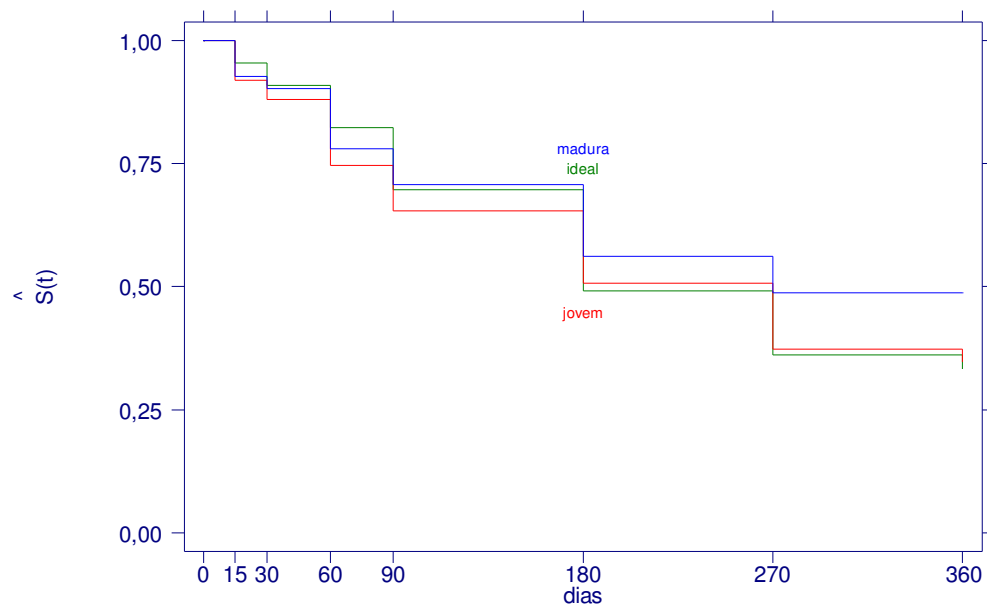


Gráfico 6 – Estimativas de Kaplan-Meier para a escolaridade materna – recordatório



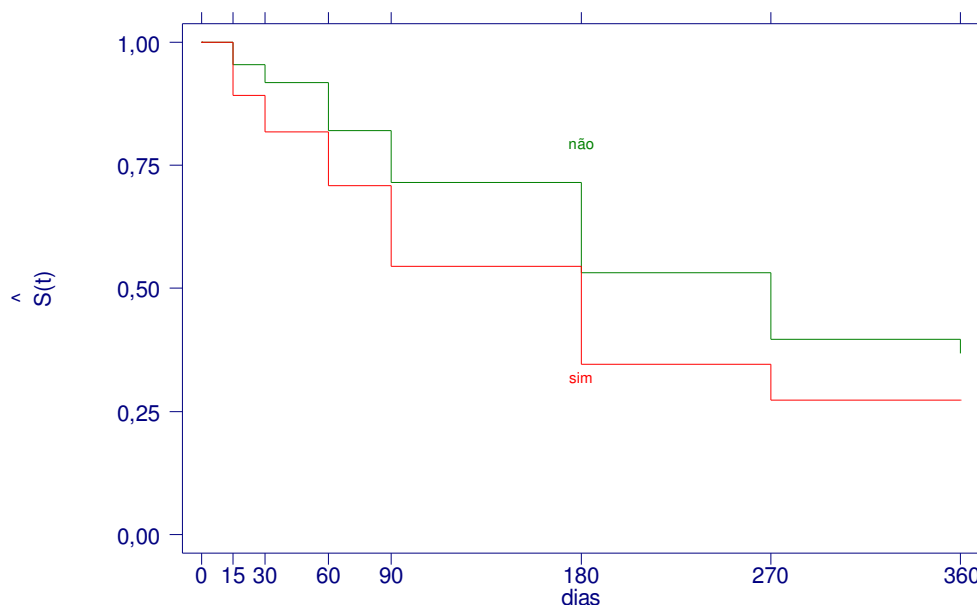
O teste de *log-rank* indica não haver diferença entre as curvas de escolaridade materna ( $p=0,0889$ ).

Gráfico 7 – Estimativas de Kaplan-Meier para a idade materna – recordatório



O teste de *log-rank* indica não haver diferença entre as curvas de idade materna ( $p=0,3490$ ).

Gráfico 8 – Estimativas de Kaplan-Meier para hábito de fumar da mãe – recordatório



O teste de *log-rank* indica haver diferença entre as curvas de idade materna ( $p=0,0240$ ).

Tabela 7 – Teste de *log-rank* – recordatório

Variáveis	% de falhas	S(360)	Log-rank
<i>sexo</i>			0,5592
masculino	33,52	0,3352	
feminino	37,08	0,3708	
<i>tipo de parto</i>			0,5328
normal	35,00	0,3500	
fórceps	32,81	0,3281	
cesárea	37,50	0,3750	
<i>raça</i>			0,6765
branca	33,18	0,3318	
não branca	38,36	0,3836	
<i>idade da mãe (anos)</i>			0,3490
até 19 anos	33,20	0,3320	
de 20 ou 34	34,67	0,3467	
34 anos e mais	48,78	0,4878	
<i>escolaridade materna</i>			0,0889
nível universitário	61,11	0,6111	
nível secundário	31,03	0,3103	
nível fundamental	36,04	0,3604	

Tabela 7 (continuação) – Teste de *log-rank* – recordatório

Variáveis	% de falhas	S(360)	<i>Log-rank</i>
<i>condição marital</i>			
casada	36,40	0,3640	0,3386
não casada	31,17	0,3117	
<i>hábito de fumar</i>			
não	36,72	0,3672	0,0240
sim	27,27	0,2727	
<i>morar com o pai</i>			
sim	35,59	0,3672	0,7552
não	33,85	0,2727	
<i>posse de bens</i>			
sim	37,04	0,3704	0,2835
não	32,64	0,3264	

O Anexo 7 apresenta a análise univariada dos modelos de Cox e complemento log-log com cada uma das variáveis. Nelas, destacamos a presença das variáveis escolaridade da mãe e hábito de fumar da mãe durante a gestação, para os dois modelos.

### 4.3 Usando as covariáveis em um modelo multivariado

As variáveis hábito de fumar e escolaridade da mãe foram selecionadas em todos os modelos univariados para entrar no modelo multivariado. O procedimento de seleção do modelo utilizado se encontra no Anexo 7.

Na Tabela 8, encontra-se o modelo para os dados do caderno diário incluindo-se todas as variáveis para se observar os valores da razão de riscos (HR). Na Tabela 9 está o resultado após o procedimento de seleção do modelo.

Tabela 8 – Estimativas dos parâmetros para todas as variáveis, segundo os modelos de Cox e complemento log-log – caderno diário

Variáveis	Total (n=360)	%	Cox – diário			Complemento log-log – diário				
			HR ajustado	IC 95%	p*	HR ajustado	IC 95%	p*		
<i>sexo</i>										
masculino	182	50,6	1				1			
feminino	178	49,4	0,9441	0,7212	1,2359	0,676	0,9248	0,7050	1,2130	0,572
<i>tipo de parto</i>										
normal	200	55,6	1				1			
fórceps	64	17,8	1,2050	0,8350	1,7389	0,319	1,2023	0,8307	1,7401	0,329
cesárea	96	26,6	0,8133	0,5827	1,1360	0,225	0,7980	0,5704	1,1165	0,188
<i>raça</i>										
branca	214	59,4	1				1			
não branca	146	40,6	1,0496	0,7916	1,3917	0,737	1,0616	0,7979	1,4123	0,682
<i>idade da mãe (anos)</i>										
entre 20 e 34	244	67,8	1				1			
20 ou menos	75	20,8	1,1331	0,8483	1,5134	0,398	1,1566	0,8635	1,5492	0,329
34 ou mais	41	11,4	1,1331	0,8483	1,5134	0,398	1,1566	0,8635	1,5492	0,329
<i>escolaridade materna</i>										
nível universitário	18	5,0	1				1			
nível secundário	145	40,3	1,5464	0,7672	3,1171	0,223	1,6205	0,8028	3,2707	0,178
nível fundamental	197	54,8	1,4785	0,7280	3,0027	0,279	1,5499	0,7633	3,1468	0,225

Tabela 8 (continuação) – Estimativas dos parâmetros para todas as variáveis, segundo os modelos de Cox e complemento log-log – caderno diário

Variáveis	Total (n=360)	%	Cox – diário			Complemento log-log – diário				
			HR ajustado	IC 95%	p*	HR ajustado	IC 95%	p*		
<i>condição marital</i>										
casada	283	78,6	1				1			
não casada	77	21,4	1,6741	0,8802	3,1843	0,116	1,8715	0,9781	3,5808	0,058
<i>hábito de fumar</i>										
não	305	84,7	1				1			
sim	55	15,3	1,4856	1,0289	2,1451	0,035	1,5086	1,0409	2,1863	0,030
<i>morar com o pai</i>										
sim	295	81,9	1				1			
não	65	18,1	0,6093	0,3013	1,2321	0,168	0,5384	0,2644	1,0963	0,080
<i>posse de bens</i>										
sim	216	60,0	1				1			
não	144	40,0	1,1226	0,8389	1,5424	0,436	1,1519	0,8584	1,5457	0,346

\* Teste de Wald

As estimativas da razão das funções de risco e seus intervalos, para todas as variáveis, são próximas entre os dois modelos aplicados ao caderno diário.

Na Tabela 9, após o procedimento de seleção, a variável que permanece em ambos os modelos é hábito de fumar da mãe. Os coeficientes são muito semelhantes, assim como os limites dados pelos intervalos de confiança.

Tabela 9 – Estimativas dos parâmetros para a variável hábito de fumar, segundo os modelos de Cox e complemento log-log – caderno diário

<i>hábito de fumar</i>	Total (n=360)	%	caderno diário							
			Cox			Complemento log-log				
			HR	IC 95%	p*	HR	IC 95%	p*		
não	305	84,7	1				1			
sim	55	15,3	1,4732	1,0297	2,1077	0,034	1,4475	1,0195	2,0552	0,039

\* Teste de Wald

Segundo o modelo de Cox, quem teve o hábito de fumar durante a gravidez apresenta um risco de deixar de dar o leite de peito 1,47 (IC95% 1,03-2,11) vezes o de quem não fumou ou, equivalentemente, o risco de deixar de dar o leite de peito de quem fumou durante a gravidez é 47% maior ao risco de quem não fumou. O modelo complemento log-log apresentou estimativas muito próximas (1,45; IC95% 1,01-2,05) ou, equivalentemente, 45% maior. Assim, nas duas modelagens para o caderno diário, os valores de HR e seus intervalos de confiança permitem dizer que o fato de a mãe fumar durante a gravidez representa um risco de desmame da criança maior do que quem não fumou.

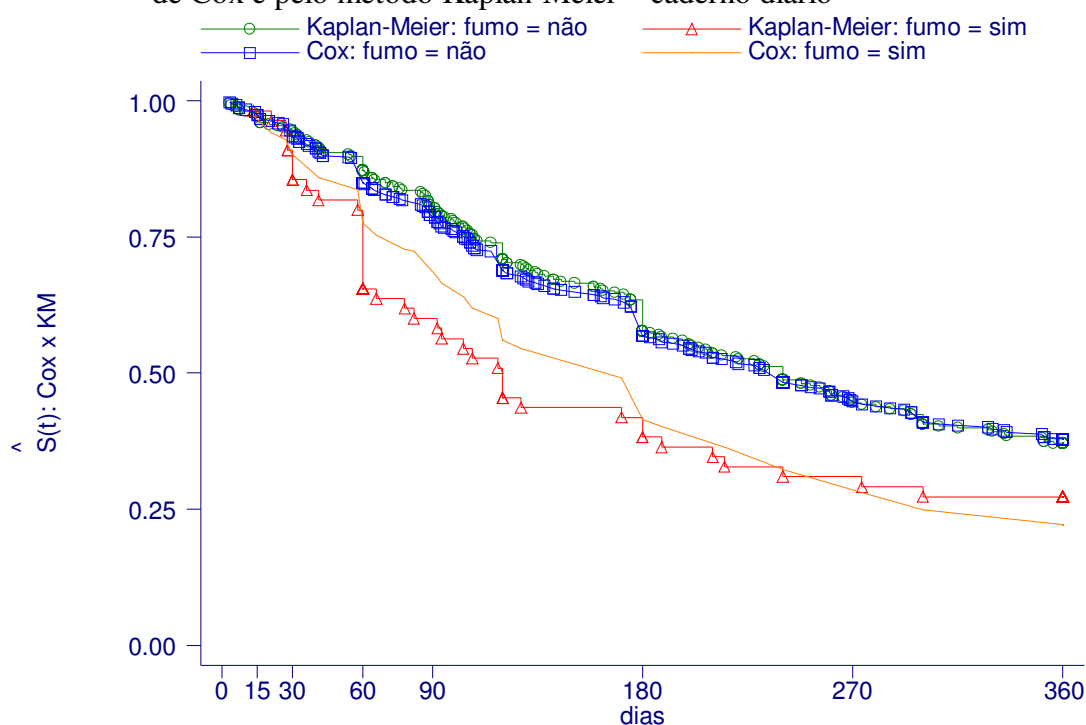
A Tabela 10 apresenta os valores estimados da função de sobrevivência, em determinados tempos, segundo os modelos complemento log-log e de Cox, segundo categorias da variável hábito de fumar. Pode ser observado que os valores nos dois modelos, segundo as categorias, praticamente se superpõem.

Tabela 10 – Estimativas da sobrevida pelos modelos segundo categorias de hábito de fumar – caderno diário

Tempo	Hábito de fumar			
	não		sim	
	Cox	Complemento log-log	Cox	Complemento log-log
0	1	1	1	1
15	0,9783	0,9705	0,9681	0,9818
30	0,9456	0,9443	0,9207	0,8545
60	0,8796	0,8718	0,8275	0,6545
90	0,7875	0,8088	0,7029	0,6000
180	0,5978	0,5772	0,4679	0,3818
270	0,4369	0,4451	0,2945	0,3091
360	0,3672	0,3170	0,2278	0,2440

O Gráfico 9 apresenta as estimativas, para o modelo de Cox do caderno diário, da variável hábito de fumar e os valores observados por meio da técnica Kaplan-Meier.

Gráfico 9 – Sobrevida, segundo categorias de hábito de fumar, estimadas pelo modelo de Cox e pelo método Kaplan-Meier – caderno diário

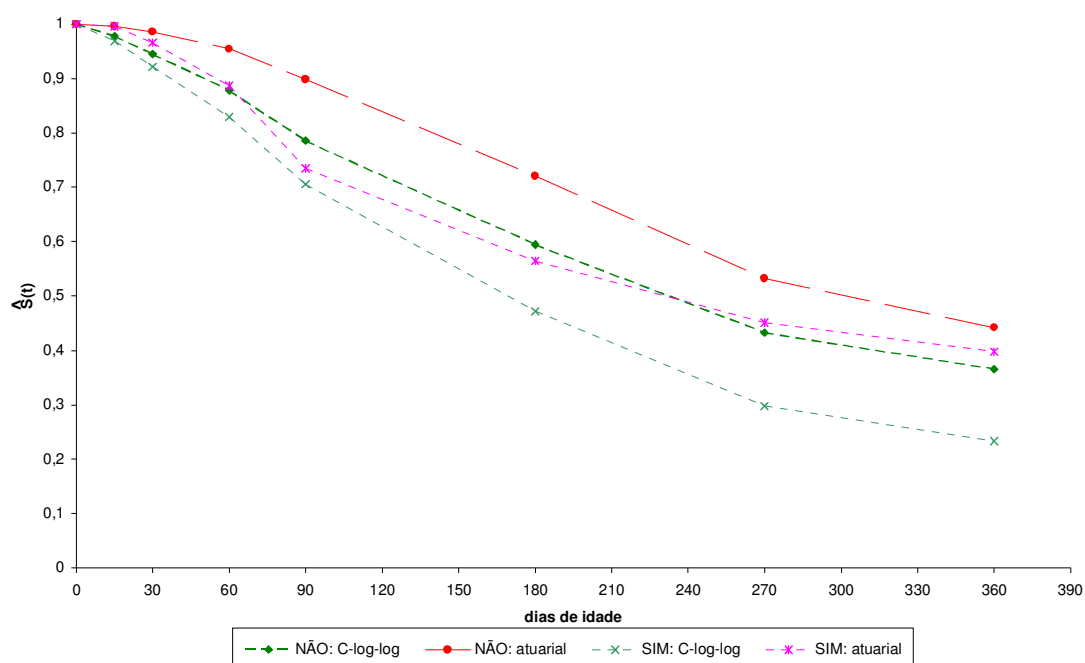


Pelo gráfico, pode ser notado que, para quem não fumou durante a gravidez, o estimador de Kaplan-Meier está próximo aos do de Cox em todos o período de tempo

de acompanhamento. De forma contrária, para aqueles que fumaram, os valores observados estão por vezes acima e por outras abaixo do valor fornecido pelo modelo. A maior diferença ocorre entre 60 e 180 dias, período em que o modelo apresenta probabilidades maiores de sobrevivência do que o observado. Isso indica que existe um problema na suposição de proporcionalidade dos riscos.

O Gráfico 10 apresenta as estimativas, para o modelo complemento log-log, da variável hábito de fumar e os valores observados.

Gráfico 10 – Sobrevida segundo categorias de hábito de fumar, estimadas pelo modelo complemento log-log e pela técnica atuarial – caderno diário



Pode ser observado no Gráfico 10 que os valores ajustados tanto para quem fumou como para quem não fumou durante a gravidez não estão muito próximas dos valores observados, mas os valores observados das duas categorias se aproximam bastante no período final.

Na Tabela 11, encontra-se o modelo incluindo-se todas as variáveis e na Tabela 12 o resultado após o procedimento de seleção do modelo, detalhado no Anexo 7, para os dados do recordatório.



Tabela 11 – Estimativas dos parâmetros para todas as variáveis, segundo os modelos de Cox e complemento log-log – recordatório

Variáveis	Cox – recordatório			Complemento log-log – recordatório				
	HR (ajustado)	IC 95%	p*	HR (ajustado)	IC 95%	p*		
<i>sexo</i>								
masculino	1			1				
feminino	0,8860	0,6805	1,1536	0,369	0,8831	0,6730	1,1418	0,329
<i>tipo de parto</i>								
normal	1			1				
fórceps	1,1929	0,8176	1,7403	0,360	1,2252	0,8390	1,7893	0,293
cesárea	1,0060	0,7333	1,3802	0,970	1,0200	0,7420	1,4021	0,903
<i>raça</i>								
branca	1			1				
não branca	0,9514	0,7208	1,2557	0,725	0,9617	0,7273	1,2717	0,784
<i>idade da mãe (anos)</i>								
entre 20 e 34	1			1				
20 ou menos	0,9893	0,6980	1,4023	0,952	0,9948	0,7016	1,4107	0,977
34 ou mais	0,8308	0,5154	1,3389	0,446	0,8062	0,4994	1,3014	0,378
<i>escolaridade materna</i>								
nível universitário	1			1				
nível secundário	2,0119	0,9115	4,4408	0,084	2,0168	0,9128	4,4541	0,083
nível fundamental	1,8515	0,8351	4,1046	0,129	1,8717	0,8436	4,1530	0,123

\* Teste de Wald

Tabela 11 (continuação) – Estimativas dos parâmetros para todas as variáveis, segundo os modelos de Cox e complemento log-log – recordatório

Variáveis	Cox – recordatório			Complemento log-log – recordatório				
	HR (ajustado)	IC 95%	p*	HR (ajustado)	IC 95%	p*		
<i>condição marital</i>								
casada	1			1				
não casada	1,6599	0,8969	3,0723	0,107	1,7166	0,9231	3,1923	0,088
<i>hábito de fumar</i>								
não	1			1				
sim	1,4298	0,9991	2,0461	0,051	1,4583	1,0159	2,0936	0,041
<i>mora com o pai</i>								
não	1			1				
sim	0,6088	0,3095	1,1978	0,151	0,5848	0,2956	1,1569	0,123
<i>posse de bens</i>								
sim	1			1				
não	1,1353	0,8512	1,5141	0,388	1,1305	0,8470	1,5090	0,405

\* Teste de Wald

As variáveis apresentaram estimativas de razão de riscos semelhantes. Após o procedimento de seleção do modelo (passos no Anexo 7), a única variável que permanece é hábito de fumar da mãe (Tabela 12).

Tabela 12 – Estimativas dos parâmetros para a variável hábito de fumar, segundo os modelos de Cox e complemento log-log – recordatório.

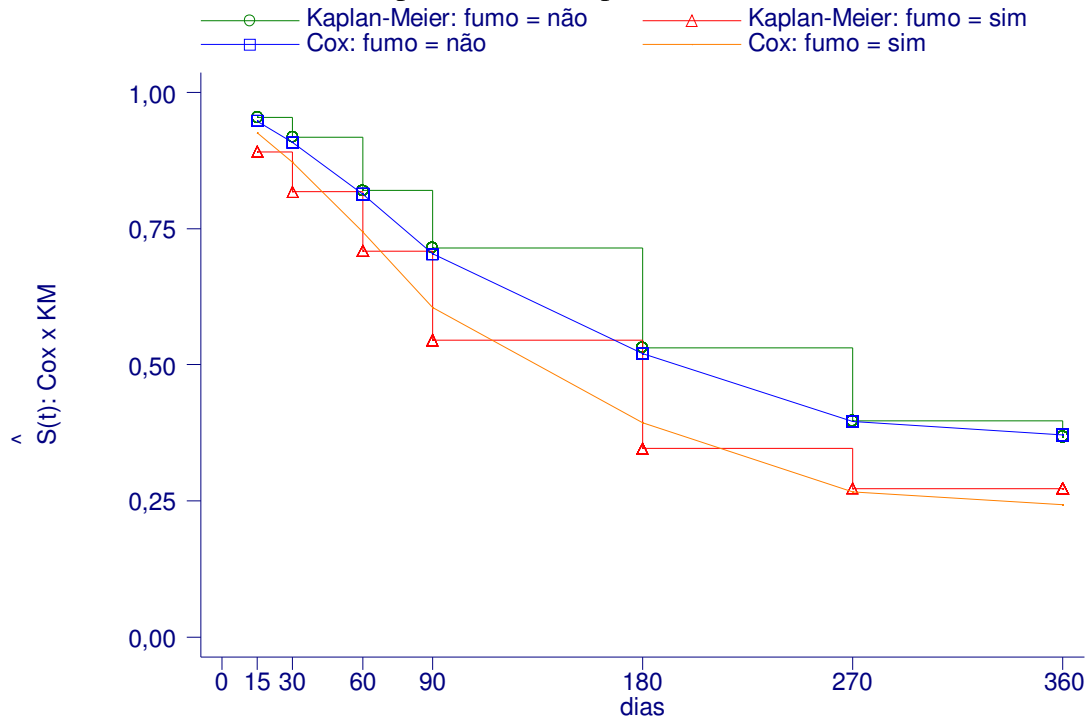
<i>hábito de fumar</i>	Total (n=360)	%	recordatório							
			Cox			Complemento log-log				
			HR	IC 95%	p*	HR	IC 95%	p*		
não	305	84,7	1				1			
sim	55	15,3	1,4594	1,0377	2,0526	0,030	1,4763	1,0134	2,0159	0,026

\* Teste de Wald

Segundo o modelo de Cox, verifica-se que quem teve o hábito de fumar durante a gravidez apresenta um risco de deixar de dar o leite de peito 1,46 (IC95% 1,04-2,05) vezes o de quem não fumou ou, equivalentemente, 46% maior. O modelo complemento log-log apresentou estimativas muito próximas (1,48, IC95% 1,01-2,02) ou, equivalentemente, 48% maior.

O Gráfico 11 apresenta as estimativas para o modelo de Cox do caderno diário, da variável hábito de fumar e os valores observados por meio da técnica Kaplan-Meier.

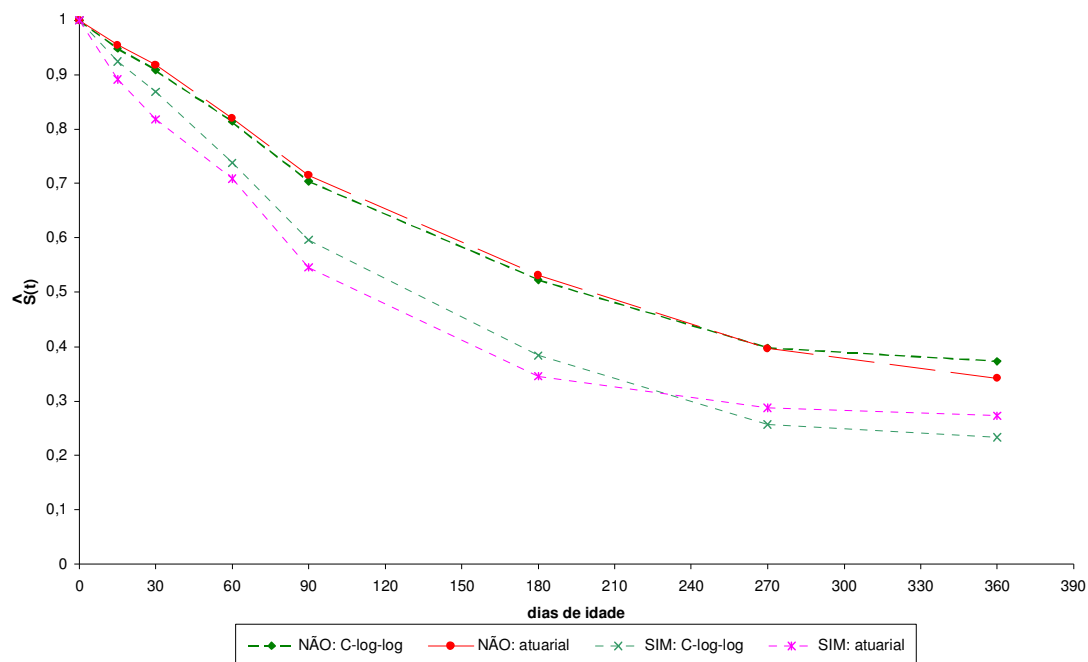
Gráfico 11 – Sobrevida, segundo categorias de hábito de fumar, estimadas pelo modelo de Cox e pelo método Kaplan-Meier – recordatório



Pelo gráfico, os valores observados das probabilidades segundo os modelos de Cox e Kaplan-Meier estão próximos em todos o período de tempo de acompanhamento. Isso indica que proporcionalidade dos riscos está satisfeita.

O Gráfico 12 apresenta os valores ajustados para  $\hat{S}(t)$  pelo modelo complemento log-log para o recordatório, além das estimativas da técnica atuarial. De forma diferente no caso do caderno diário, os riscos não fogem à proporcionalidade ao longo do tempo.

Gráfico 12 – Sobrevida segundo categorias de hábito de fumar, estimada pelo modelo complemento log-log e atuarial – recordatório



#### 4.3.1 Comentários

As covariáveis utilizadas nos modelos, todas categóricas por natureza ou categorizadas, foram definidas de forma que o valor zero indicasse a categoria de referência. Ambos os modelos permitem a interpretação dos parâmetros por meio da razão de riscos (HR).

O ajuste do modelo de Cox aos dados do caderno diário é direta, bastando definir os tempos de desmame, a variável que indica a falha/censura e quais covariáveis utilizar.

Para se ajustar o modelo complemento log-log, os dados do caderno diário tiveram que ser agrupados em intervalos de tempo. Portanto, o tempo original é alterado e isso é justamente o que ocorre quando se trabalha com a aplicação de um recordatório em que a informação é obtida em intervalos. Para se ajustar o modelo é necessário replicar os indivíduos e suas covariáveis, no banco de dados. O número de linhas a mais no banco de dados corresponde ao número de intervalos pelos quais o indivíduo passou.

Os resultados dos dois modelos, aplicados à informação do caderno diário, apresentam resultados bastante próximos. Ambos indicam a variável hábito de fumar da mãe como importante para se compreender o desmame. Os valores das estimativas da razão de riscos e dos intervalos de confiança também são bastante próximos.

O modelo complemento log-log, aplicado aos dados do recordatório, é adequado para os dados de uma pesquisa de coorte retrospectiva. O tempo fornecido nesse tipo de pesquisa é basicamente de origem intervalar.

O ajuste do modelo complemento log-log aos dados do caderno diário serve para avaliar, com uma mesma técnica, se os dados fornecidos pelos recordatórios possuem semelhança aos do caderno diário. Os resultados da Tabela 11 mostram que as estimativas são praticamente iguais entre os modelos, e a seleção posterior dos modelos finais aponta para a mesma variável. Além desse ajuste, o modelo de Cox, mais apropriado ao caderno diário, apresentou resultados semelhantes aos do modelo complemento log-log (Tabela 8).

## 5 Discussão

Devido ao fato de cada um dos cadernos conter 60 dias para anotação, existe uma tendência de registro do evento em múltiplos de 60. A anotação dos eventos pode estar sendo completada pela mãe no momento da entrevista e implica um viés do entrevistador.

O viés de recordação, ou de memória, pode estar presente em um estudo de abordagem retrospectiva. Aparentemente, não deve existir diferença na resposta das mães para os diversos níveis das variáveis. A comparação entre as duas abordagens, utilizando uma mesma técnica, pode apontar diferenças que podem estar representando esse viés. Partindo da suposição que a abordagem prospectiva é a mais próxima da medida real, as diferenças que surgem entre esta e os recordatórios podem ser devidas à precisão das medidas e/ou a vieses, sendo o mais importante deles, o de memória.

A idéia de se trabalhar inicialmente sem as covariáveis foi a de analisar o efeito somente das informações dos dois tipos de estudo no tempo de amamentação. Cada tipo de estudo apresenta vantagens e desvantagens, sendo que é importante notar que associado ao estudo de coorte retrospectiva está o fator de memória da mãe. Nessa análise, nota-se que os estimadores apresentam resultados próximos ao do Kaplan-Meier, indicando que os dois recordatórios forneceram resultados numericamente concordantes com o do caderno diário e semelhantes entre si o que reforça que o recordatório pode ser confiável para avaliar o tempo de amamentação.

Deve-se lembrar que as censuras foram dadas por motivo de fim de estudo e não de perda e, na ausência de censuras, o estimador atuarial é idêntico ao do modelo complemento log-log. No último intervalo do recordatório, quando ocorrem as censuras, o modelo complemento log-log fornece estimativa mais alta que a da técnica atuarial.

A técnica atuarial é bastante simples e não pressupõe riscos proporcionais mas não tem estrutura para se medir o efeito de mais de uma covariável, enquanto que o modelo complemento log-log permite a inclusão de diversas covariáveis no modelo.

Existem pequenas diferenças entre as estimativas para os recordatórios e para o caderno diário. As diferenças encontradas entre as medidas calculadas para os recordatórios e o caderno diário podem ser frutos de falha de memória da mãe. A imprecisão da informação obtida é inerente a um estudo de coorte retrospectiva. Apesar das diferenças, pode-se ter uma boa idéia do desmame no decorrer do tempo por meio dos recordatórios.

As quedas bruscas na função de sobrevivência pelo Kaplan-Meier mostram que as entrevistas domiciliares têm um papel ainda mais importante a cumprir para se detectar e corrigir falhas na pesquisa. Para se tentar melhorar o preenchimento do caderno diário, as visitas domiciliares podem ser mais frequentes. Porém, o aumento de visitas, além de encarecer uma pesquisa, pode aumentar a influência da pesquisa no preenchimento da amamentação.

Quando são consideradas as covariáveis, foram aplicados os modelos de Cox e complemento log-log para as informações do caderno diário. As estimativas dos efeitos de cada covariável são bem parecidas entre os modelos. As estimativas fornecem uma noção de quais variáveis podem ser importantes na amamentação. Isso não significa que devem ser utilizados estudos de coorte retrospectiva para se medir os efeitos de covariáveis no lugar de estudos de coorte, mas eles podem ser úteis para ajudar a identificar covariáveis importantes. Deve-se notar, ainda, que a construção de uma coorte retrospectiva não é fácil, pois depende da identificação e localização de todas as mães em todo o período de estudo. A falta de informações pode produzir um estudo que mais se aproxima de um estudo de corte transversal.

Nessa análise, a falta de proporcionalidade dos riscos ocorre para os dois modelos. Caso fossem utilizados modelos que comportem variáveis tempo-dependentes, os resultados para os dois modelos poderiam ser diferentes mas, com a semelhança aqui apresentada, não se acredita que houvesse muita alteração.

Ao se utilizar o modelo de Cox e complemento log-log aos dados do recordatório, as estimativas do risco de falhas foram bastante semelhantes entre si. Assim, os modelos apresentaram resultados semelhantes tanto com o dado retrospectivo como o do prospectivo.



Observando-se apenas o modelo complemento log-log, os dados do caderno diário apresentaram estimativas muito parecidas aos do recordatório indicando que as duas fontes de informação apresentam certa concordância.

A mesma semelhança entre as estimativas é encontrada quando se utiliza o modelo de Cox para as duas fontes.

Em geral, as estimativas das funções de sobrevida fornecidas pelos dois modelos finais são bastante parecidas. Mais ainda, somente uma e a mesma variável permaneceu no modelo final para todas as técnicas, com estimativas muito próximas, reforçando semelhança entre as várias abordagens. Esta variável, hábito de fumar da mãe, aparece como fator de risco em diversos artigos normalmente acompanhadas de outros fatores (Levy 2002; Gigante 2000; Najdawi 1999; Rutishauser 1992).

## **6 Conclusão**

O modelo de Cox apresentou estimativas semelhantes tanto para a coorte retrospectiva quanto para a coorte prospectiva.

O modelo complemento log-log apresentou estimativas semelhantes tanto para a coorte retrospectiva quanto para a coorte prospectiva.

Os resultados do modelo de Cox e modelo complemento log-log apresentaram estimativas semelhantes para os dados da coorte retrospectiva.

Os resultados do modelo de Cox e modelo complemento log-log apresentaram estimativas semelhantes para os dados da coorte prospectiva.

O risco de uma criança deixar mais precocemente o leite de peito naquelas com mães que fumaram durante a gravidez é maior comparado ao risco para crianças com mães que não fumaram.

## 7 Referências

- Abada TSJ, Trovato F, Lalu N. Determinants of breastfeeding in the Philippines: a survival analysis. *Social Science and Medicine* 2001, 52: 71-81.
- Assis AMO, Prado MS, Freitas MC, Silva RCR, Ramos LB, Machado AD. Prática do aleitamento materno em comunidades rurais do semi-árido baiano. *Revista de Saúde Pública* 1994, 28(5): 380-4.
- Breslow NE. A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. *Biometrika* 1970; 57: 579-594.
- Breslow NE. Covariance analysis of censored survival data. *Biometrics* 1974; 30: 89-100.
- Carstensen B. Regression models for interval censored data: application to HIV infection in Danish homosexual men. *Statistics in Medicine* 1996; 15: 2177-2189.
- Chalita LVAS, Colossimo EA, Demétrio CGB. *Likelihood approximation and discrete models for interval-censored data under Cox regression model*. UNESP NOV 1999. Relatório Técnico 99-01.
- Collett D. *Modelling Survival Data in Medical Research*. London: Chapman Hill; 1994.
- Corrente JE, Chalita LVAS, Moreira JA. Choosing between Cox proportional hazards and logistic models for interval-censored data via bootstrap. *Journal of Applied Statistics* JAN 2003, 30 (1): 37-47
- Cox DR. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society* 1972, 34 (Series B): 187-200.
- Cox DR, Oakes D. *Analysis of Survival Data*. London: Chapman and Hall, 1984.
- Escamilla RP, Lutter CK, Wickham C et al. Identification of risk factors for short breastfeeding in Mexico City through survival analysis. *Ecol Food and Nutrition* 1997, 36(1): 43-64.
- Farrington CP. Interval censored survival data: a generalized linear model approach. *Statistics in Medicine* 1996; 15: 283-292.
- Gehan EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrics* 1965; 52: 203-223.

Gigante DP, Victora CG, Barros FC. Nutrição materna e duração da amamentação em uma coorte de nascimento de Pelotas, RS. *Revista de Saúde Pública* 2000, 34(3): 259-65.

Hardy EE et al. A prática de amamentação no município de Paulínia, Estado de São Paulo, Brasil. *Revista de Saúde Pública* 1982, 16: 337-45.

Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. *Statistics in Medicine* 1995, 14: 1707-1723.

Hitchcock NE et al. The growth of breast fed and artificially fed infants from birth to twelve months. *Acta Paediatrica Scandinavica* 1985; 74: 240-5.

Hosmer DW, Lemeshow S. *Applied Survival Analysis: regression modeling of time to event data*. New York: John Wiley and Sons; 1999.

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of Chronic Diseases* 1955; 1: 12-32.

Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. New York: Wiley; 1980.

Klein JP, Moeschberger ML. *Survival Analysis: techniques for censored and truncated data*. New York: Springer; 1997.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research*. New York: Van Nostrand Reinhold; 1982.

Kosterman R, Hawkins JD, Guo J, Catalano RF, Abbott RD. The dynamics of alcohol and marijuana initiation: Patterns and predictors of first use in adolescence. *American Journal of Public Health* MAR 2000, 90 (3): 360-366.

Kummer SC, Giugliani ERJ, Susin LO et al. Evolução do padrão de aleitamento materno. *Revista de Saúde Pública* 2000, 34(2): 143-8;

Lancelot R, Lesnoff M, McDermott JJ. Use of Akaike information criteria for model selection and inference. An application to assess prevention of gastrointestinal parasitism and respiratory mortality of Guinean goats in Kolda, Senegal. *Preventive Veterinary Medicine* NOV 15 2002, 55 (4): 217-240.

Lawless JF. *Statistical models and methods for lifetime data*. New York: John Wiley & Sons, Inc.; 1982.

Lee ET. *Statistical methods for survival data analysis*. New York: John Wiley & Sons, Inc.; 1992.

Levy SM, Slager SL, Warren JJ et al. Association of pacifier use, digit sucking, and child care attendance with cessation of breastfeeding. *Journal of Family and Practice* MAY 2002, 51(5): U15-U19.

Lilienfeld AM. *Foundations of epidemiology*. New York: Oxford University Press; 1976.

Comissão Nacional de Ética em Pesquisa. *Normas para pesquisa envolvendo seres humanos (Res. CNS 196/96 e outras) = Rules on research involving human subjects / Conselho Nacional de Saúde*. Brasília: Ministério da Saúde; 2000.

McCullagh P, Nelder JA. *Generalized linear models*. 2<sup>nd</sup>. ed. London: Chapman Hill; 1989.

OPAS/OMS. *Indicadores para evaluar las practicas de lactancia maternal*. OMS/CED/SER/91. 14. Genebra; 1991.

Passos MC, Lamounier JA, Silva CAM et al. Práticas de amamentação no município de Ouro Preto, MG, Brasil. *Revista de Saúde Pública* 2000; 34(6): 617-22.

Pereira MG. *Epidemiologia: teoria e prática*. Rio de Janeiro: Guanabara Koogan; 1999.

Peto R, Peto J. Asymptotically efficient rank invariance test procedures (with discussion). *Journal of the American Statistical Association* 1972; 72: 854-858.

Prentice RL, Gloecker LA. Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 1978, 34: 57-67.

Radrigan M et al. *Alimentación en el primer año de vida*. Ver. Chil. Pediatr. 1989, 60 (Suppl): 9-19.

Rea MF, Venâncio SI, Batista LE, Santos RG, Greiner T. Possibilidades e limitações da amamentação entre mulheres trabalhadoras formais. *Revista de Saúde Pública* 1997, 31(2): 149-56.

Rothman KJ, Greenland S. *Modern Epidemiology*. 2<sup>nd</sup>. ed. Philadelphia: Lippincott – Raven; 1998.

Rutishauser IHE, Carlin JB. Body mass index and duration of breast feeding: a survival analysis during the first six months of life. *Journal of Epidemiology and Community Health* 1992, 46: 559-565.

Scott JA, Aitkin I, Binns CW et al. Factors associated with the duration of breastfeeding among women in Perth, Australia. *Acta Paediatric* APR 1999, 88(4): 416-421.

- Szklo M, Nieto FJ. *Epidemiology: beyond the basics*. Maryland: Aspen; 2000.
- STATA CORP. Stata statistical software: release 7.0. College Station (TX): Stata Corporation; 2001.
- Therneau TM, Grambsch PM. *Modeling survival data: extending the Cox model*. New York: Springer; 2000.
- Weiderpass E, Barros FC, Victora CG, Tomasi E, Halpern R. Incidência e duração da amamentação conforme o tipo de parto: estudo longitudinal no sul do Brasil. *Revista de Saúde Pública* 1998; 32(3): 225-31.

## **Anexos**

## **Anexo 1 – Termo de consentimento**



**TERMO DE CONSENTIMENTO PARA PARTICIPAÇÃO NA PESQUISA  
ALIMENTAÇÃO NO PRIMEIRO ANO DE VIDA**

(de acordo com a Resolução 196 de 10/10/1996 do Conselho Nacional de Saúde)

Eu, \_\_\_\_\_, livremente, permito que meu(minha) filho(a) participe da pesquisa “**Alimentação no primeiro ano de vida**” sob responsabilidade dos pesquisadores José Maria Pacheco de Souza, Professor Titular da Faculdade de Saúde Pública da USP, Sonia Buongiorno de Souza, Professora Doutora da Faculdade de Saúde Pública da USP, Suely Godoy Agostinho Gimeno, Professora Doutora da Universidade Federal de São Paulo-Escola Paulista de Medicina- e Arnaldo Augusto Franco de Siqueira, Professor Titular da Faculdade de Saúde Pública da USP.

**Fui informado(a) sobre o seguinte:**

**Objetivo da Pesquisa:** Verificar a prática alimentar no primeiro ano de vida e sua relação com o crescimento e o desenvolvimento da criança.

**Participação:** Responder às perguntas necessárias e permitir que sejam feitas medidas de peso, de comprimento, de perímetros cefálico e torácico, a cada visita domiciliar, ao longo do primeiro ano de vida.

**Riscos:** Não haverá riscos para a integridade física, mental ou moral de meu(minha) filho(a).

**Benefícios:** As informações obtidas nesta pesquisa poderão ser úteis cientificamente e de ajuda para todos.

**Privacidade:** Os dados individualizados serão confidenciais. Os resultados coletivos serão divulgados nos meios científicos.

**Contatos com os pesquisadores:** Terei acesso a telefone para esclarecimento de dúvidas ou reclamações.

**Desistência:** Poderei desistir a qualquer momento da minha participação, sem qualquer consequência para mim.

---

Assinatura

**Anexo 2 – Formulário 1**

**FORMULÁRIO 1  
IDENTIFICAÇÃO**

1. NOME DA MÃE	
2. NOME DA CRIANÇA	
3. ENDEREÇO	
4. N° DO FORMULÁRIO	_1_
5. N° DA CRIANÇA	_ _ _
6. DATA	_ _ _ _ _
7. ENTREVISTADORA	_ _ _
8. N° DO REGISTRO HOSPITALAR	_ _ _ _ _ _
9. RELAÇÃO DA MÃE COM A USP Servidor não docente(0) Servidor docente(1) Dep. de não docente(2) Dep. de docente(3) Aluno(4) Não USP(5)	_ _
10. RELAÇÃO DO PAI COM A USP Servidor não docente(0) Servidor docente(1) Dep. de não docente(2) Dep. de docente(3) Aluno(4) Não USP(5)	_ _

**CARACTERÍSTICAS DA CRIANÇA**

11. DATA DE NASCIMENTO	_ _ _ _ _
12. SEXO Masculino(0) Feminino(1)	_ _
13. COR/RAÇA/ETNIA Branca(0) Preta(1) Parda/Mulata(2) Amarela/Oriental(3)	_ _
14. ORDEM DE NASCIMENTO (nascidos vivos)	_ _ _
15. PESO AO NASCER (g)	_ _ _ _ _
16. COMPRIMENTO AO NASCER (cm)	_ _ _
17. PERÍMETRO CEFÁLICO (cm)	_ _ _
18. PERÍMETRO TORÁCICO (cm)	_ _ _

**CONDIÇÕES DA GESTAÇÃO E PARTO**

|\_|\_|\_|\_|

19. TEMPO DE GESTAÇÃO (semanas)	_ _ _ _
20. GANHO DE PESO DA MÃE NA GESTAÇÃO (kg)	_ _ _ _
21. NÚMERO DE CONSULTAS NO PRÉ-NATAL	_ _ _ _
22. PARTO Normal(0) Fórceps(1) Cesárea(2)	_ _
23. NÚMERO DE GESTAÇÕES ANTERIORES	_ _ _ _
24. HÁBITO DE FUMAR DA MÃE ANTES DA GESTAÇÃO Não(00) Número de cigarros por dia( )	_ _ _ _
25. HÁBITO DE FUMAR DA MÃE DURANTE A GESTAÇÃO Não(00) Número de cigarros por dia( )	_ _ _ _
26. HÁBITO DE BEBER DA MÃE ANTES DA GESTAÇÃO Não(0) Ocasionalmente(1) Semanalmente(2) Diariamente (3)	_ _
27. HÁBITO DE BEBER DA MÃE DURANTE A GESTAÇÃO Não (0) Ocasionalmente (1) Semanalmente (2) Diariamente (3)	_ _

**CARACTERÍSTICAS FAMILIARES**

28. CONDIÇÃO MARITAL DA MÃE Solteira(0) Casada(1) Viúva(2) Separada(3)	__
29. MÃE MORA COM PAI DA CRIANÇA Não(0) Sim(1)	__
30. IDADE DA MÃE	__ __
31. IDADE DO PAI	__ __
32. LOCAL DE NASCIMENTO DA MÃE_____ S.Paulo(0) Grande S.Paulo(1) Est.S.Paulo(2) Outro Estado(3) Outro país(4)	__
33. LOCAL DE NASCIMENTO DO PAI_____ S.Paulo(0) Grande S.Paulo(1) Est.S.Paulo(2) Outro Estado(3) Outro país(4)	__
34. ESCOLARIDADE DA MÃE Grau( ) Série( )	__ __
35. ESCOLARIDADE DO PAI Grau( ) Série( )	__ __
36. OCUPAÇÃO DA MÃE	__
37. OCUPAÇÃO DO PAI	__

	_ _ _ _ _
38. RENDA TOTAL MENSAL DA FAMÍLIA (R\$)	_ _ _ _ _ _ _ _ _ _
39. N° DE CRIANÇAS RESIDENTES NA CASA	_ _ _ _
40. N° DE CRIANÇAS QUE VÃO À CRECHE OU BERÇÁRIO	_ _ _ _
41. NÚMERO DE CRIANÇAS QUE VÃO À ESCOLA	_ _ _ _
42. NÚMERO DE CRIANÇAS QUE FICAM EM OUTROS LOCAIS, EM CONTATO COM OUTRAS CRIANÇAS (VIZINHO, PARENTE, IGREJA). LOCAL: _____	_ _ _ _
43. QUEM ESCOLHEU O NOME DA CRIANÇA Mãe(0) Pai(1) Juntos(2) Outra(3)	_ _ _ _
44. NÚMERO DE FUMANTES RESIDENTES NA CASA	_ _ _ _
45. O PAI FUMA? Não(00) N° Cigarros p/dia( )	_ _ _ _
46. HÁBITO DE BEBER DO PAI Não(0) Ocasionalmente (1) Semanalmente(2) Diariamente (3)	_ _ _ _
47. NÚMERO DE IRMÃOS VIVOS DA CRIANÇA	_ _ _ _
48. IDADE DO IRMÃO MAIS VELHO, EM ANOS INTEIROS	_ _ _ _
49. AMAMENTOU SÓ NO PEITO ATÉ ( ) MESES	_ _ _ _
50. AMAMENTOU NO PEITO ATÉ ( ) MESES	_ _ _ _
51. QUAL A RAZÃO DO DESMAME DO PEITO	_ _ _ _
52. IDADE DO IRMÃO MAIS NOVO, EM ANOS INTEIROS	_ _ _ _
53. AMAMENTOU SÓ NO PEITO ATÉ ( ) MESES	_ _ _ _
54. AMAMENTOU NO PEITO ATÉ ( ) MESES	_ _ _ _
55. QUAL A RAZÃO DO DESMAME DO PEITO	_ _ _ _
56. DIABETES NA MÃE Não(0) Sim, tipo I(1) Sim, tipo II(2) Não sabe(9)	_ _ _ _
57. DIABETES NO PAI Não(0) Sim, tipo I(1) Sim, tipo II(2) Não sabe(9)	_ _ _ _
58. DIABETES EM IRMÃO Não(0) Sim, tipo I(1) Sim, tipo II(2) Não sabe(9)	_ _ _ _

**Anexo 3 – Formulário 2**

**FORMULÁRIO 2**

1. NOME DA MÃE	
2. NOME DA CRIANÇA	
4. N° DO FORMULÁRIO	2
5. N° DA CRIANÇA	__ __ __
6. DATA	__ __ __
7. ENTREVISTADORA	__ __
8. N° DO REGISTRO HOSPITALAR	__ __ __ __ __ __

**CONDIÇÕES DO DOMICÍLIO**

10. CONSTRUÇÃO EM ALVENARIA Não (0) Sim (1)	__
11. NÚMERO DE CÔMODOS (salas e quartos)	__ __
12. NÚMERO DE BANHEIROS (incl. lavabo e empreg.)	__
13. NÚMERO DE MORADORES	__ __
14. FAVELA Não (0) Sim (1)	__
15. ÁGUA ENCANADA Não tem(0) Dentro de casa(1)Fora de casa(2)	__
16. ESGOTO Não (0) Sim (1)	__
17. PAVIMENTAÇÃO Não (0) Sim (1)	__
18. ELETRICIDADE Não (0) Sim (1)	__
19. HABITAÇÃO COLETIVA Não (0) Sim (1)	__

**POSSE DE BENS**

20. NÚMERO DE COMPUTADORES	__
21. NÚMERO DE CARROS	__
22. NÚMERO DE APARELHOS DE TV	
23. TEM TV PAGA (NET, TVA)? Não (0) Sim (1)	__
24. NÚMERO DE TELEFONES/CELULARES (LINHAS)	__
25. NÚMERO DE RÁDIOS/SOM	__
26. NÚMERO DE VIDEO-CASSETES	__
27. TEM MÁQ.DE LAVAR ROUPAS? Não (0) Sim (1)	__
28. NÚMERO DE GELADEIRAS/FREEZERS	__
29. TEM MÁQ.DE LAVAR LOUÇA? Não (0) Sim (1)	__

**Anexo 4 – Formulário 3**



**FORMULÁRIO 3**

1. NOME DA MÃE	
2. NOME DA CRIANÇA	
4. N° DO FORMULÁRIO	_3_
5. N° DA CRIANÇA	_ _ _ _
6. DATA	_ _ _ _ _ _ _
7. ENTREVISTA	_
8. ENTREVISTADORA	_ _ _
9. N° DO REGISTRO HOSPITALAR	_ _ _ _ _ _ _
10. IDADE DA CRIANÇA (em dias)	_ _ _ _
11. HÁBITO DE FUMAR Não(00) N° de cigarros/dia( )	_ _ _
12. HÁBITO DE BEBER DA MÃE Não(0) Ocasionalmente(3) Semanalmente(2) Diariamente(1)	_ _

**CRESCIMENTO / CUIDADOS DE SAÚDE**

13. PESO (gramas)	_ _ _ _ _ _ _
14. COMPRIMENTO (cm)	_ _ _ _
15. PERÍMETRO CEFÁLICO (cm)	_ _ _ _
16. PERÍMETRO TORÁCICO (cm)	_ _ _ _
17. VACINA BCG Não(000) Sim(DIAS DE IDADE)	_ _ _ _
18. VACINA DPT Não(000) Sim(DIAS DE IDADE)	_ _ _ _
19. VACINA SABIN Não(000) Sim(DIAS DE IDADE)	_ _ _ _
20. VACINA SARAMPO Não(000) Sim(DIAS DE IDADE)	_ _ _ _
21. HÁ LUGAR DEFINIDO PARA O BEBÊ QUANDO SE ALIMENTA? Não(0) Sim(1)Qual:	_ _
22. EXISTE UM LUGAR DEFINIDO PARA O BEBÊ REPOUSAR Não(0) Sim(1) Qual:	_ _
23. EXISTE UM LUGAR DEFINIDO PARA A HIGIENE DO BEBÊ Não(0) Sim(1) Qual:	_ _
24. EXISTE UM LUGAR DEFINIDO PARA O BEBÊ NO CONVÍVIO DA FAMÍLIA Não(0) Sim(1) Qual:	_ _
25. RESPONSÁVEL PELA ALIMENTAÇÃO DA CRIANÇA Mãe(0) Outro Parente(1) Creche(3) Outros(4) Quem?	_ _

**CONSULTAS MÉDICAS**

|\_|\_|\_|

26. NÚMERO DE CONSULTAS MÉDICAS DE ROTINA  
(Locais, médicos, endereços, orientações)

|\_|\_|\_|

27. NÚMERO DE CONSULTAS MÉDICAS POR DOENÇA  
(Locais, médicos, internações, endereços,  
razões, orientações)

|\_|\_|\_|

**DESMAME**    |\_\_|\_\_|\_\_|

28. CONTINUA DANDO SÓ LEITE DE PEITO Não (0) Sim (1)	__
29. CONTINUA DANDO LEITE DE PEITO Não (0) Sim (1)	__
30. QUAL ALIMENTO INTRODUZIU PRIMEIRO?	__
COMEÇOU A DAR OUTRO(S) ALIMENTO(S) PORQUE	
31. O LEITE ERA FRACO Não (0) Sim (1)	__
32. NÃO TINHA LEITE            Não (0) Sim (1)	__
33. O LEITE SECOU Não (0) Sim (1)	__
34. O BEBÊ SENTIA FOME Não (0) Sim (1)	__
35. O BEBÊ REJEITAVA O PEITO Não (0) Sim (1)	__
36. O BEBÊ FICOU DOENTE Não (0) Sim (1) Dizer a doença..... Quem diagnosticou?	__
37. O PEITO FICOU MACHUCADO Não (0) Sim (1)	__
38. TINHA DOR AO AMAMENTAR Não (0) Sim (1)	__
39. A MÃE FICOU DOENTE Não (0) Sim (1) Dizer a doença..... Quem diagnosticou?.....	__
40. PRECISOU TRABALHAR Não (0) Sim (1)	__
41. COMEÇOU A TRABALHAR Não (0) Sim (1)	__
42. VOLTOU A TRABALHAR Não (0) Sim (1)	__
43. O BEBÊ FOI PARA A CRECHE Não (0) Sim (1)	__
44. PRECISOU DEIXAR O BEBÊ COM ALGUÉM Não (0) Sim (1) Quem?	__
45. ALGUÉM ORIENTOU A INTRODUÇÃO? Conta própria (0) Médico (1) Outra (3) Quem?	__
46. SEDE/CALOR	__
47. CÓLICA/PRISÃO DE VENTRE	__



## **Anexo 5 – Questionário recordatório**

**“Questionário recordatório”**

**RECORDATÓRIO** |\_\_|\_\_|\_\_|

55. DEU LEITE DE PEITO (0) NÃO (1) SIM	__
56. ATÉ QUANDO A CRIANÇA RECEBEU LEITE DE PEITO Continua (0)	<b>Meses</b>  __ __  <b>Dias</b>  __ __ __
57. POR QUE DEIXOU DE DAR LEITE DE PEITO	__ __
58. ATÉ QUANDO DEU SÓ LEITE DE PEITO Continua (0)	<b>Meses</b>  __ __  <b>Dias</b>  __ __ __
59. QUANDO FOI INTRODUZIDO CHÁ, ÁGUA E/OU SUCO	<b>Meses</b>  __ __  <b>Dias</b>  __ __ __
60. QUANDO FOI INTRODUZIDO OUTRO TIPO DE LEITE	<b>Meses</b>  __ __  <b>Dias</b>  __ __ __
61. QUANDO FOI INTRODUZIDO O PRIMEIRO ALIMENTO NÃO LÍQUIDO (papinha, fruta, mingau, etc.) Qual?_____	<b>Meses</b>  __ __  <b>Dias</b>  __ __ __
62. PESO DA CRIANÇA AO NASCER (gramas)	__ __ __ __
63. PESO DA CRIANÇA AOS 6 MESES (gramas)	__ __ __ __
64. COMPRIMENTO DA CRIANÇA AO NASCER (cm)	__ __
65. COMPRIMENTO DA CRIANÇA AOS 6 MESES (cm)	__ __

**CONDIÇÕES SÓCIO-ECONÔMICAS**

64. RENDA FAMILIAR	__ __ __ __
65. OCUPAÇÃO DA MÃE	__ __
66. OCUPAÇÃO DO PAI	__ __

**Anexo 6 – Caderno diário**





## **Anexo 7 – Escolha dos modelos**

### Modelo: Cox; fonte de informação: caderno diário

#### MODELO COM TODAS AS VARIÁVEIS

```
Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1222,9722
Iteration 2: log likelihood = -1222,7589
Iteration 3: log likelihood = -1222,7586
Refining estimates:
Iteration 0: log likelihood = -1222,7586
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
LR chi2(13) = 17,62
Prob > chi2 = 0,1725
```

	_t					
	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
sexo		,9259365	,1278752	-0,56	0,577	,7063618 1,213767
_Iparto_1		1,283143	,2473283	1,29	0,196	,8794358 1,872172
_Iparto_2		,8522068	,144533	-0,94	0,346	,6111989 1,188249
racag		1,034325	,149491	0,23	0,815	,7791712 1,373034
_Icidademg~1		1,016522	,1868422	0,09	0,929	,7090243 1,457378
_Icidademg~2		,8502513	,2056429	-0,67	0,502	,5292668 1,365903
_Icescolma~1		1,463618	,5314179	1,05	0,294	,7184043 2,981853
_Icescolma~2		1,422859	,5216563	0,96	0,336	,6935693 2,919
_Imaritalg_2		1,767998	,5578044	1,81	0,071	,9526366 3,281228
fumodurg		1,611194	,2997911	2,56	0,010	1,118836 2,320218
cmorapai		,5666372	,1992354	-1,62	0,106	,2844553 1,128746
_Icbensgr_1		1,176431	,1766076	1,08	0,279	,8765624 1,578884

#### MODELO COM SEXO

```
Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1231,503
Iteration 2: log likelihood = -1231,503
Refining estimates:
Iteration 0: log likelihood = -1231,503
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
LR chi2(1) = 0,13
Prob > chi2 = 0,7176
```

	_t					
	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
sexo		,9529871	,1269014	-0,36	0,718	,7340735 1,237185

#### MODELO COM TIPO DE PARTO

```
Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1229,9424
Iteration 2: log likelihood = -1229,9337
Iteration 3: log likelihood = -1229,9337
Refining estimates:
Iteration 0: log likelihood = -1229,9337
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
LR chi2(2) = 3,27
Prob > chi2 = 0,1950
```

	_t					
	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_Iparto_1		1,203373	,2099579	1,06	0,289	,8548445 1,693999
_Iparto_2		,8310714	,1364403	-1,13	0,260	,6024114 1,146525

**MODELO COM RAÇA**

```

Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1231,5614
Iteration 2: log likelihood = -1231,5614
Refining estimates:
Iteration 0: log likelihood = -1231,5614
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
Log likelihood = -1231,5614 LR chi2(1) = 0,01
Prob > chi2 = 0,9057
-----
      _t |
      _d | Haz. Ratio Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
      racag | ,9840155 ,1339866 -0,12 0,906 ,7535286 1,285003
-----

```

**MODELO COM IDADE DA MÃE**

```

Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1230,6482
Iteration 2: log likelihood = -1230,6399
Iteration 3: log likelihood = -1230,6399
Refining estimates:
Iteration 0: log likelihood = -1230,6399
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
Log likelihood = -1230,6399 LR chi2(2) = 1,86
Prob > chi2 = 0,3951
-----
      _t |
      _d | Haz. Ratio Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
      _Icidademg~1 | 1,06042 ,1763535 0,35 0,724 ,7654534 1,469053
      _Icidademg~2 | ,7622752 ,1736575 -1,19 0,233 ,4877481 1,191319
-----

```

**MODELO COM ESCOLARIDADE DA MÃE**

```

Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1230,0799
Iteration 2: log likelihood = -1230,0228
Iteration 3: log likelihood = -1230,0226
Refining estimates:
Iteration 0: log likelihood = -1230,0226
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
Log likelihood = -1230,0226 LR chi2(2) = 3,09
Prob > chi2 = 0,2131
-----
      _t |
      _d | Haz. Ratio Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
      _Icescolma~1 | 1,709641 ,5971586 1,54 0,125 ,8621639 3,39016
      _Icescolma~2 | 1,743181 ,6021443 1,61 0,108 ,885754 3,430613
-----

```

A variável ESCOLARIDADE DA MÃE deve ser selecionada.

**MODELO COM CONDIÇÃO MARITAL DA MÃE**

Iteration 0: log likelihood = -1231,5684  
 Iteration 1: log likelihood = -1231,2937  
 Iteration 2: log likelihood = -1231,2933  
 Refining estimates:  
 Iteration 0: log likelihood = -1231,2933  
 Cox regression -- Breslow method for ties  
 No. of subjects = 360 Number of obs = 360  
 No. of failures = 226  
 Time at risk = 76797  
 LR chi2(1) = 0,55  
 Prob > chi2 = 0,4582  
 Log likelihood = -1231,2933

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_lmaritalg_2			1,12884	,1823032	0,75	0,453	,8225588 1,549167

**MODELO COM HÁBITO DE FUMAR DA MÃE**

Iteration 0: log likelihood = -1231,5684  
 Iteration 1: log likelihood = -1228,7519  
 Iteration 2: log likelihood = -1228,6491  
 Iteration 3: log likelihood = -1228,649  
 Refining estimates:  
 Iteration 0: log likelihood = -1228,649  
 Cox regression -- Breslow method for ties  
 No. of subjects = 360 Number of obs = 360  
 No. of failures = 226  
 Time at risk = 76797  
 LR chi2(1) = 5,84  
 Prob > chi2 = 0,0157  
 Log likelihood = -1228,649

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_lfumodurg_1			1,556805	,2717156	2,54	<b>0,011</b>	1,105785 2,191785

A variável HÁBITO DE FUMAR deve ser selecionada.

**MODELO COM MORAR O PAI DA CRIANÇA COM A MÃE**

Iteration 0: log likelihood = -1231,5684  
 Iteration 1: log likelihood = -1231,568  
 Iteration 2: log likelihood = -1231,568  
 Refining estimates:  
 Iteration 0: log likelihood = -1231,568  
 Cox regression -- Breslow method for ties  
 No. of subjects = 360 Number of obs = 360  
 No. of failures = 226  
 Time at risk = 76797  
 LR chi2(1) = 0,00  
 Prob > chi2 = 0,9774  
 Log likelihood = -1231,568

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
cmorapai			1,005004	,1769399	0,03	0,977	,711715 1,419153

**MODELO COM POSSE DE BENS**

```

Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1230,7376
Iteration 2: log likelihood = -1230,7373
Refining estimates:
Iteration 0: log likelihood = -1230,7373
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
Log likelihood = -1230,7373 LR chi2(1) = 1,66
Prob > chi2 = 0,1973

```

_____t	_____d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_____+	_____+					
_Icbensgr_1		1,190612	,1603209	1,30	0,195	,9144339 1,550201

**MODELO COM AS VARIÁVEIS SELECIONADAS - ESCOLARIDADE DA MÃE E HÁBITO DE FUMAR**

```

Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1227,6362
Iteration 2: log likelihood = -1227,4877
Iteration 3: log likelihood = -1227,4876
Refining estimates:
Iteration 0: log likelihood = -1227,4876
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
Log likelihood = -1227,4876 LR chi2(3) = 8,16
Prob > chi2 = 0,0428

```

_____t	_____d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_____+	_____+					
_Icescolma~1		1,621438	,5680024	1,38	0,168	,8160519 3,221683
_Icescolma~2		1,622756	,5635895	1,39	0,163	,8215383 3,205375
fumodurg		1,511777	,2655672	2,35	0,019	1,071423 2,133116

**MODELO COM HÁBITO DE FUMAR**

```

Iteration 0: log likelihood = -1231,5684
Iteration 1: log likelihood = -1228,7519
Iteration 2: log likelihood = -1228,6491
Iteration 3: log likelihood = -1228,649
Refining estimates:
Iteration 0: log likelihood = -1228,649
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 226
Time at risk = 76797
Log likelihood = -1228,649 LR chi2(1) = 5,84
Prob > chi2 = 0,0157

```

_____t	_____d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_____+	_____+					
fumodurg		1,556805	,2717156	2,54	0,011	1,105785 2,191785

O TESTE DA RAZÃO DE VEROSSIMILHANÇAS, A PARTIR DOS DOIS MODELOS ANTERIORES, INDICA A SAÍDA DA VARIÁVEL ESCOLARIDADE DA MÃE.

```

chi2(2) = 2,32
Prob > chi2 = 0,3131

```

DESSA FORMA, O MODELO FINAL FOI O QUE CONTÉM SOMENTE HÁBITO DE FUMAR DA MÃE.

## Modelo: complemento log-log; fonte de informação: caderno diário

### MODELO COM TODAS AS VARIÁVEIS

i.parto                \_Iparto\_0-2               (naturally coded; \_Iparto\_0 omitted)  
i.cidademg1           \_Icidademg1\_0-2           (naturally coded; \_Icidademg1\_0 omitted)  
i.cescolmatg          \_Icescolmat\_0-2          (naturally coded; \_Icescolmat\_0 omitted)  
i.maritalg            \_Imaritalg\_1-2            (naturally coded; \_Imaritalg\_1 omitted)  
i.cbensgr             \_Icbensgr\_0-1             (naturally coded; \_Icbensgr\_0 omitted)  
Iteration 0:   log likelihood = -618,76296  
Iteration 1:   log likelihood = -615,91883  
Iteration 2:   log likelihood = -615,91553  
Iteration 3:   log likelihood = -615,91553  
Complementary log-log regression

Number of obs        =        1953  
Zero outcomes        =        1727  
Nonzero outcomes    =        226

Log likelihood = -615,91553

Wald chi2(19)        =        814,96  
Prob > chi2         =        0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-4,096632	,4754118	-8,62	0,000	-5,028422 -3,164841
int2	-3,649701	,4383328	-8,33	0,000	-4,508818 -2,790585
int3	-2,774447	,3936146	-7,05	0,000	-3,545917 -2,002977
int4	-3,068349	,4113169	-7,46	0,000	-3,874515 -2,262182
int5	-1,543602	,3664606	-4,21	0,000	-2,261851 -,8253521
int6	-1,851692	,3795491	-4,88	0,000	-2,595595 -1,10779
int7	-2,193583	,4057435	-5,41	0,000	-2,988826 -1,39834
sexo	-,1033342	,1368808	-0,75	0,450	-,3716156 ,1649472
_Iparto_1	,2578886	,1936224	1,33	0,183	-,1216044 ,6373816
_Iparto_2	-,1374107	,1698039	-0,81	0,418	-,4702202 ,1953989
racag	,0257182	,1449491	0,18	0,859	-,2583768 ,3098131
_Icidademg~1	,0259236	,1842609	0,14	0,888	-,3352211 ,3870684
_Icidademg~2	-,1431811	,2415463	-0,59	0,553	-,6166031 ,3302409
_Icescolma~1	,4017985	,3624004	1,11	0,268	-,3084932 1,11209
_Icescolma~2	,3731954	,3656085	1,02	0,307	-,3433841 1,089775
_Imaritalg_2	,5918108	,3185242	1,86	0,063	-,0324853 1,216107
fumodurg	,4554219	,1844659	2,47	0,014	,0938755 ,8169684
cmorapai	-,6297439	,3541825	-1,78	0,075	-1,323929 ,064441
_Icbensgr_1	,2000892	,1509637	1,33	0,185	-,0957941 ,4959725

### MODELO COM SEXO

Iteration 0:   log likelihood = -626,80656  
Iteration 1:   log likelihood = -624,55677  
Iteration 2:   log likelihood = -624,55262  
Iteration 3:   log likelihood = -624,55262

Complementary log-log regression                Number of obs        =        1953  
Zero outcomes        =        1727  
Nonzero outcomes    =        226

Log likelihood = -624,55262

Wald chi2(8)        =        828,41  
Prob > chi2         =        0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,541396	,3226239	-10,98	0,000	-4,173727 -2,909065
int2	-3,100047	,2659711	-11,66	0,000	-3,62134 -2,578753
int3	-2,238147	,1855004	-12,07	0,000	-2,601722 -1,874573
int4	-2,543357	,222615	-11,42	0,000	-2,979674 -2,107039
int5	-1,033664	,1284949	-8,04	0,000	-1,28551 -,7818188
int6	-1,358507	,1677962	-8,10	0,000	-1,687382 -1,029633
int7	-1,703347	,2196382	-7,76	0,000	-2,13383 -1,272864
sexo	-,0577914	,133481	-0,43	0,665	-,3194093 ,2038266

**MODELO COM TIPO DE PARTO**

Iteration 0: log likelihood = -625,50415  
 Iteration 1: log likelihood = -623,06839  
 Iteration 2: log likelihood = -623,06463  
 Iteration 3: log likelihood = -623,06463  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -623,06463  
 Wald chi2(9) = 826,14  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,562436	,321732	-11,07	0,000	-4,193019 -2,931853
int2	-3,11887	,2647002	-11,78	0,000	-3,637673 -2,600067
int3	-2,255987	,1833948	-12,30	0,000	-2,615435 -1,89654
int4	-2,561054	,2205733	-11,61	0,000	-2,99337 -2,128739
int5	-1,052264	,1248175	-8,43	0,000	-1,296902 -,807626
int6	-1,376832	,1649613	-8,35	0,000	-1,700151 -1,053514
int7	-1,716642	,21801	-7,87	0,000	-2,143934 -1,28935
_Iparto_1	,190503	,1749099	1,09	0,276	-,1523141 ,5333202
_Iparto_2	-,1757055	,1645259	-1,07	0,286	-,4981702 ,1467593

**MODELO COM RAÇA**

Iteration 0: log likelihood = -626,89311  
 Iteration 1: log likelihood = -624,64426  
 Iteration 2: log likelihood = -624,64011  
 Iteration 3: log likelihood = -624,64011  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226  
 Wald chi2(8) = 828,55  
 Prob > chi2 = 0,0000

Log likelihood = -624,64011

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,563317	,3209284	-11,10	0,000	-4,192325 -2,934308
int2	-3,122022	,2637409	-11,84	0,000	-3,638944 -2,605099
int3	-2,260272	,1819995	-12,42	0,000	-2,616985 -1,90356
int4	-2,565462	,2197541	-11,67	0,000	-2,996172 -2,134752
int5	-1,055578	,1238228	-8,52	0,000	-1,298266 -,81289
int6	-1,380914	,1640308	-8,42	0,000	-1,702409 -1,05942
int7	-1,727098	,2159017	-8,00	0,000	-2,150257 -1,303938
racag	-,0152844	,1364828	-0,11	0,911	-,2827858 ,252217

**MODELO COM IDADE DA MÃE**

Iteration 0: log likelihood = -625,98904  
 Iteration 1: log likelihood = -623,68655  
 Iteration 2: log likelihood = -623,68248  
 Iteration 3: log likelihood = -623,68248  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226  
 Wald chi2(9) = 827,00  
 Prob > chi2 = 0,0000

Log likelihood = -623,68248

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,555755	,3193654	-11,13	0,000	-4,181699 -2,92981
int2	-3,113789	,2618052	-11,89	0,000	-3,626918 -2,60066
int3	-2,25253	,1793844	-12,56	0,000	-2,604117 -1,900943
int4	-2,556895	,2176083	-11,75	0,000	-2,9834 -2,130391
int5	-1,04536	,1200676	-8,71	0,000	-1,280688 -,8100321
int6	-1,367548	,1613235	-8,48	0,000	-1,683736 -1,05136
int7	-1,708545	,2137355	-7,99	0,000	-2,127459 -1,289631
_Icidademg~1	,067016	,1667172	0,40	0,688	-,2597437 ,3937757
_Icidademg~2	-,272748	,2281387	-1,20	0,232	-,7198916 ,1743956

**MODELO COM ESCOLARIDADE DA MÃE**

Iteration 0: log likelihood = -625,39273  
 Iteration 1: log likelihood = -623,09003  
 Iteration 2: log likelihood = -623,08588  
 Iteration 3: log likelihood = -623,08588  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -623,08588  
 Wald chi2(9) = 825,91  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-4,097919	,4574155	-8,96	0,000	-4,994437 -3,201401
int2	-3,655624	,4191015	-8,72	0,000	-4,477048 -2,8342
int3	-2,792721	,3727838	-7,49	0,000	-3,523364 -2,062078
int4	-3,096127	,3918683	-7,90	0,000	-3,864174 -2,328079
int5	-1,582051	,3453158	-4,58	0,000	-2,258858 -,9052445
int6	-1,902503	,3593481	-5,29	0,000	-2,606813 -1,198194
int7	-2,248988	,3856996	-5,83	0,000	-3,004945 -1,493031
_Icescolma~1	,5378555	,3496178	1,54	<b>0,124</b>	-,1473828 1,223094
_Icescolma~2	,5592154	,345735	1,62	<b>0,106</b>	-,1184128 1,236844

A variável ESCOLARIDADE DA MÃE deve ser selecionada.

**MODELO COM CONDIÇÃO MARITAL DA MÃE**

Iteration 0: log likelihood = -626,65663  
 Iteration 1: log likelihood = -624,39813  
 Iteration 2: log likelihood = -624,39398  
 Iteration 3: log likelihood = -624,39398  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -624,39398  
 Wald chi2(8) = 828,18  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,595365	,3184778	-11,29	0,000	-4,21957 -2,97116
int2	-3,153842	,2609326	-12,09	0,000	-3,665261 -2,642423
int3	-2,291832	,1781032	-12,87	0,000	-2,640908 -1,942756
int4	-2,596382	,2163271	-12,00	0,000	-3,020376 -2,172389
int5	-1,085738	,1170678	-9,27	0,000	-1,315187 -,8562894
int6	-1,409735	,1582016	-8,91	0,000	-1,719804 -1,099665
int7	-1,754846	,2111363	-8,31	0,000	-2,168665 -1,341026
_Imaritalg_2	,1163361	,1619503	0,72	0,473	-,2010807 ,4337529

**MODELO COM HÁBITO DE FUMAR DA MÃE**

Iteration 0: log likelihood = -624,20787  
 Iteration 1: log likelihood = -621,70772  
 Iteration 2: log likelihood = -621,70397  
 Iteration 3: log likelihood = -621,70397  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -621,70397  
 Wald chi2(8) = 824,38  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,652182	,3185916	-11,46	0,000	-4,27661 -3,027754
int2	-3,20756	,2608947	-12,29	0,000	-3,718904 -2,696216
int3	-2,336596	,1772878	-13,18	0,000	-2,684074 -1,989118
int4	-2,636525	,2154507	-12,24	0,000	-3,058801 -2,21425
int5	-1,123037	,1153724	-9,73	0,000	-1,349163 -,8969117
int6	-1,448251	,1573743	-9,20	0,000	-1,756699 -1,139803
int7	-1,799919	,2111242	-8,53	0,000	-2,213715 -1,386123
fumodurg	,445765	,1751304	2,55	<b>0,011</b>	,1025156 ,7890143

A variável HÁBITO DE FUMAR deve ser selecionada.



**MODELO COM MORAR O PAI DA CRIANÇA COM A MÃE**

Iteration 0: log likelihood = -626,89798  
 Iteration 1: log likelihood = -624,65055  
 Iteration 2: log likelihood = -624,64639  
 Iteration 3: log likelihood = -624,64639  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -624,64639  
 Wald chi2(8) = 828,56  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1	-3,569403	,3178086	-11,23	0,000	-4,192297	-2,94651
int2	-3,127999	,2600931	-12,03	0,000	-3,637772	-2,618226
int3	-2,266158	,1768462	-12,81	0,000	-2,61277	-1,919546
int4	-2,571371	,2153985	-11,94	0,000	-2,993544	-2,149197
int5	-1,061536	,1157793	-9,17	0,000	-1,288459	-,8346129
int6	-1,386987	,1576325	-8,80	0,000	-1,695941	-1,078033
int7	-1,733223	,2108147	-8,22	0,000	-2,146413	-1,320034
cmorapai	-,0003533	,1764757	-0,00	0,998	-,3462393	,3455327

**MODELO COM POSSE DE BENS**

Iteration 0: log likelihood = -625,87363  
 Iteration 1: log likelihood = -623,60206  
 Iteration 2: log likelihood = -623,59789  
 Iteration 3: log likelihood = -623,59789  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -623,59789  
 Wald chi2(8) = 826,95  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1	-3,653079	,3220056	-11,34	0,000	-4,284199	-3,02196
int2	-3,211677	,2652527	-12,11	0,000	-3,731562	-2,691791
int3	-2,34765	,1839316	-12,76	0,000	-2,70815	-1,987151
int4	-2,651859	,2211162	-11,99	0,000	-3,085239	-2,21848
int5	-1,140189	,1254388	-9,09	0,000	-1,386045	-,8943335
int6	-1,461794	,1640154	-8,91	0,000	-1,783258	-1,14033
int7	-1,809426	,2160175	-8,38	0,000	-2,232812	-1,38604
_Icbensgr_1	,1965932	,1349994	1,46	<b>0,145</b>	-,0680008	,4611871

A variável POSSE DE BENS deve ser selecionada.

**MODELO COM AS VARIÁVEIS SELECIONADAS - ESCOLARIDADE DA MÃE, HÁBITO DE FUMAR E POSSE DE BENS**

Iteration 0: log likelihood = -622,51522  
 Iteration 1: log likelihood = -619,96156  
 Iteration 2: log likelihood = -619,95782  
 Iteration 3: log likelihood = -619,95782  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -619,95782  
 Wald chi2(11) = 821,51  
 Prob > chi2 = 0,0000

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-4,115862	,4576238	-8,99	0,000	-5,012788	-3,218936
int2		-3,671124	,4192995	-8,76	0,000	-4,492936	-2,849312
int3		-2,798192	,3728669	-7,50	0,000	-3,528998	-2,067387
int4		-3,096384	,3918832	-7,90	0,000	-3,864461	-2,328307
int5		-1,579211	,3452992	-4,57	0,000	-2,255985	-,9024371
int6		-1,897784	,359301	-5,28	0,000	-2,602001	-1,193567
int7		-2,249237	,3857162	-5,83	0,000	-3,005227	-1,493247
_Icescolma~1		,4468205	,3525151	1,27	0,205	-,2440964	1,137737
_Icescolma~2		,4011505	,3574412	1,12	0,262	-,2994215	1,101722
fumodurg		,4108512	,1763511	2,33	0,020	,0652095	,7564929
_Icbensgr_1		,1560069	,1449432	1,08	0,282	-,1280765	,4400904

**MODELO COM ESCOLARIDADE DA MÃE, HÁBITO DE FUMAR**

Iteration 0: log likelihood = -623,07639  
 Iteration 1: log likelihood = -620,53806  
 Iteration 2: log likelihood = -620,53431  
 Iteration 3: log likelihood = -620,53431

Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -620,53431  
 Wald chi2(10) = 822,36  
 Prob > chi2 = 0,0000

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-4,113318	,4575969	-8,99	0,000	-5,010191	-3,216444
int2		-3,668193	,4192628	-8,75	0,000	-4,489933	-2,846453
int3		-2,796869	,3728463	-7,50	0,000	-3,527634	-2,066103
int4		-3,095585	,3918719	-7,90	0,000	-3,86364	-2,32753
int5		-1,578731	,3452878	-4,57	0,000	-2,255483	-,901979
int6		-1,899645	,359318	-5,29	0,000	-2,603895	-1,195395
int7		-2,251012	,3857328	-5,84	0,000	-3,007034	-1,494989
_Icescolma~1		,4836013	,3506733	1,38	0,168	-,2037058	1,170908
_Icescolma~2		,487584	,3475983	1,40	0,161	-,1936961	1,168864
fumodurg		,4158454	,1762243	2,36	0,018	,0704522	,7612386

O TESTE DA RAZÃO DE VEROSSIMILHANÇAS, A PARTIR DOS DOIS MODELOS ANTERIORES (UM COM ESCOLARIDADE DA MÃE, HÁBITO DE FUMAR E POSSE DE BENS; E O OUTRO MODELO COM ESCOLARIDADE DA MÃE E HÁBITO DE FUMAR), INDICA A SAÍDA DA VARIÁVEL POSSE DE BENS.

chi2(1) = 1,15  
 Prob > chi2 = 0,2829

**MODELO COM HÁBITO DE FUMAR**

Iteration 0: log likelihood = -624,20787  
 Iteration 1: log likelihood = -621,70772  
 Iteration 2: log likelihood = -621,70397  
 Iteration 3: log likelihood = -621,70397  
 Complementary log-log regression

Number of obs = 1953  
 Zero outcomes = 1727  
 Nonzero outcomes = 226

Log likelihood = -621,70397  
 Wald chi2(8) = 824,38  
 Prob > chi2 = 0,0000

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
int1	-3,652182	,3185916	-11,46	0,000	-4,27661 -3,027754
int2	-3,20756	,2608947	-12,29	0,000	-3,718904 -2,696216
int3	-2,336596	,1772878	-13,18	0,000	-2,684074 -1,989118
int4	-2,636525	,2154507	-12,24	0,000	-3,058801 -2,21425
int5	-1,123037	,1153724	-9,73	0,000	-1,349163 -,8969117
int6	-1,448251	,1573743	-9,20	0,000	-1,756699 -1,139803
int7	-1,799919	,2111242	-8,53	0,000	-2,213715 -1,386123
fumodurg	,445765	,1751304	2,55	0,011	,1025156 ,7890143

O TESTE DA RAZÃO DE VEROSSIMILHANÇAS, A PARTIR DOS DOIS MODELOS ANTERIORES (UM COM ESCOLARIDADE DA MÃE E HÁBITO DE FUMAR; E O OUTRO MODELO COM HÁBITO DE FUMAR), INDICA A SAÍDA DA VARIÁVEL ESCOLARIDADE DA MÃE.

chi2(1) = 2,34  
 Prob > chi2 = 0,3105

DESSA FORMA, O MODELO FINAL É O QUE CONTÉM SOMENTE HÁBITO DE FUMAR DA MÃE.

## Modelo: Cox; fonte de informação: recordatório de 12 meses

### MODELO COM TODAS AS VARIÁVEIS

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1273.5606
Iteration 2: log likelihood = -1273.2973
Iteration 3: log likelihood = -1273.2959
Iteration 4: log likelihood = -1273.2959
Refining estimates:
Iteration 0: log likelihood = -1273.2959
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(12) = 14.55
Log likelihood = -1273.2959 Prob > chi2 = 0.2669

```

	_t					
	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
sexo		.8860258	.1192917	-0.90	0.369	.6805238 1.153585
_Iparto_1		1.192889	.2298785	0.92	0.360	.8176499 1.740335
_Iparto_2		1.00604	.1622979	0.04	0.970	.7333251 1.380175
racag		.951421	.1347212	-0.35	0.725	.7208462 1.255749
_Icidademg~1		.9893397	.176091	-0.06	0.952	.6979774 1.402328
_Icidademg~2		.8307634	.2023092	-0.76	0.446	.5154554 1.338948
_Icescolma~1		2.011883	.8127281	1.73	0.084	.9114823 4.440761
_Icescolma~2		1.851472	.7520614	1.52	0.129	.835145 4.104613
_Imaritalg_2		1.659957	.5213905	1.61	0.107	.896882 3.072261
fumodurg		1.429787	.2614623	1.96	0.051	.999111 2.046109
cmorapai		.6088527	.2101938	-1.44	0.151	.3094942 1.197766
_Icbensgr_1		1.135271	.1668093	0.86	0.388	.8511961 1.514153

### MODELO COM SEXO

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1280.4348
Iteration 2: log likelihood = -1280.4348
Refining estimates:
Iteration 0: log likelihood = -1280.4348
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 0.27
Log likelihood = -1280.4348 Prob > chi2 = 0.6008

```

	_t					
	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
sexo		.9337052	.1224438	-0.52	0.601	.72208 1.207353

### MODELO COM TIPO DE PARTO

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1280.1472
Iteration 2: log likelihood = -1280.1461
Iteration 3: log likelihood = -1280.1461
Refining estimates:
Iteration 0: log likelihood = -1280.1461
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(2) = 0.85
Log likelihood = -1280.1461 Prob > chi2 = 0.6533

```

	_t					
	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_Iparto_1		1.153427	.2029597	0.81	0.417	.8169784 1.628431
_Iparto_2		.966035	.1507995	-0.22	0.825	.7114085 1.311797

**MODELO COM RAÇA**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1280.4677
Iteration 2: log likelihood = -1280.4677
Refining estimates:
Iteration 0: log likelihood = -1280.4677
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 0.21
Log likelihood = -1280.4677 Prob > chi2 = 0.6483

```

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
	racag		.9406012	.1265788	-0.46	0.649	.7225326 1.224485

**MODELO COM IDADE DA MÃE**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1279.6045
Iteration 2: log likelihood = -1279.5933
Iteration 3: log likelihood = -1279.5933
Refining estimates:
Iteration 0: log likelihood = -1279.5933
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(2) = 1.96
Log likelihood = -1279.5933 Prob > chi2 = 0.3759

```

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
	_Icidademg~1		1.001699	.1632297	0.01	0.992	.7278327 1.378615
	_Icidademg~2		.7359382	.1707269	-1.32	0.186	.4670625 1.159599

**MODELO COM ESCOLARIDADE DA MÃE**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1278.0039
Iteration 2: log likelihood = -1277.8325
Iteration 3: log likelihood = -1277.8311
Iteration 4: log likelihood = -1277.8311
Refining estimates:
Iteration 0: log likelihood = -1277.8311
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(2) = 5.48
Log likelihood = -1277.8311 Prob > chi2 = 0.0645

```

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
	_Icescolma~1		2.248308	.8795221	1.07	<b>0.088</b>	1.066612 4.839938
	_Icescolma~2		2.12979	.8273998	1.15	<b>0.092</b>	.8946133 4.56057

A variável ESCOLARIDADE DA MÃE deve ser selecionada.

**MODELO COM CONDIÇÃO MARITAL DA MÃE**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1280.1541
Iteration 2: log likelihood = -1280.1533
Refining estimates:
Iteration 0: log likelihood = -1280.1533
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 0.84
Log likelihood = -1280.1533 Prob > chi2 = 0.3603

```

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_lmaritalg_2			1.155946	.1806946	0.93	0.354	.8509025 1.570345

**MODELO COM HÁBITO DE FUMAR DA MÃE**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1278.4533
Iteration 2: log likelihood = -1278.4005
Iteration 3: log likelihood = -1278.4005
Refining estimates:
Iteration 0: log likelihood = -1278.4005
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 4.34
Log likelihood = -1278.4005 Prob > chi2 = 0.0372

```

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
fumodurg			1.459445	.2539385	2.17	<b>0.030</b>	1.037723 2.052551

A variável HÁBITO DE FUMAR deve ser selecionada.

**MODELO COM MORAR O PAI DA CRIANÇA COM A MÃE**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1280.5266
Iteration 2: log likelihood = -1280.5265
Refining estimates:
Iteration 0: log likelihood = -1280.5265
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 0.09
Log likelihood = -1280.5265 Prob > chi2 = 0.7638

```

	_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
cmorapai			1.052357	.1777246	0.30	0.763	.7558057 1.465265

**MODELO COM POSSE DE BENS**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1280.0084
Iteration 2: log likelihood = -1280.0082
Refining estimates:
Iteration 0: log likelihood = -1280.0082
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 1.13
Log likelihood = -1280.0082 Prob > chi2 = 0.2884

```

_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____
_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____
_Icbensgr_1		1.152266	.1531723	1.07	0.286	.8879759 1.495216

**MODELO COM AS VARIÁVEIS SELECIONADAS - ESCOLARIDADE DA MÃE E HÁBITO DE FUMAR**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1276.2152
Iteration 2: log likelihood = -1276.0045
Iteration 3: log likelihood = -1276.0033
Iteration 4: log likelihood = -1276.0033
Refining estimates:
Iteration 0: log likelihood = -1276.0033
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(3) = 9.14
Log likelihood = -1276.0033 Prob > chi2 = 0.0275

```

_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____
_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____
_Icescolma~1		2.154242	.8444502	1.06	0.080	.9221447 4.64473
_Icescolma~2		2.005879	.7823896	0.98	0.094	.9038884 4.308384
fumodurg		1.415257	.2476319	1.98	0.047	1.00438 1.994218

**MODELO COM HÁBITO DE FUMAR**

```

Iteration 0: log likelihood = -1280.5717
Iteration 1: log likelihood = -1278.4533
Iteration 2: log likelihood = -1278.4005
Iteration 3: log likelihood = -1278.4005
Refining estimates:
Iteration 0: log likelihood = -1278.4005
Cox regression -- Breslow method for ties
No. of subjects = 360 Number of obs = 360
No. of failures = 233
Time at risk = 75390
LR chi2(1) = 4.34
Log likelihood = -1278.4005 Prob > chi2 = 0.0372

```

_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____
_t	_d	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____	_____+_____
fumodurg		1.459445	.2539385	2.17	0.030	1.037723 2.052551

O TESTE DA RAZÃO DE VEROSSIMILHANÇAS, A PARTIR DOS DOIS MODELOS ANTERIORES, INDICA A SAÍDA DA VARIÁVEL ESCOLARIDADE DA MÃE.

```

chi2(2) = 4,79
Prob > chi2 = 0,0910

```

DESSA FORMA, O MODELO FINAL FOI O QUE CONTÉM SOMENTE HÁBITO DE FUMAR DA MÃE.

## Modelo: complemento log-log; fonte de informação: recordatório de 12 meses

**MODELO COM TODAS AS VARIÁVEIS**

Iteration 0: log likelihood = -643.47575  
 Iteration 1: log likelihood = -641.33537  
 Iteration 2: log likelihood = -641.33109  
 Iteration 3: log likelihood = -641.33109  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233

Log likelihood = -641.33109  
 Wald chi2(19) = 827.39  
 Prob > chi2 = 0.0000

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-3.581534	.4563244	-7.85	0.000	-4.475914	-2.687155
int2		-3.813361	.4737243	-8.05	0.000	-4.741843	-2.884878
int3		-2.84949	.4298639	-6.63	0.000	-3.692007	-2.006972
int4		-2.574228	.4248166	-6.06	0.000	-3.406853	-1.741603
int5		-1.836242	.4114628	-4.46	0.000	-2.642694	-1.029789
int6		-1.917258	.4188643	-4.58	0.000	-2.738217	-1.096299
int7		-3.339939	.51435	-6.49	0.000	-4.348046	-2.331831
sexo		-.1317099	.1348716	-0.98	0.329	-.3960534	.1326336
_Iparto_1		.2031216	.1932203	1.05	0.293	-.1755832	.5818265
_Iparto_2		.0197828	.1623617	0.12	0.903	-.2984403	.338006
racag		-.0390183	.1425543	-0.27	0.784	-.3184196	.240383
_Icidademg~1		-.0051703	.1781842	-0.03	0.977	-.354405	.3440643
_Icidademg~2		-.2154283	.2443146	-0.88	0.378	-.6942762	.2634196
_Icescolma~1		.7015329	.4044555	1.73	0.083	-.0911854	1.494251
_Icescolma~2		.6268465	.4066282	1.54	0.123	-.1701301	1.423823
_Imaritalg_2		.5403683	.3165276	1.71	0.088	-.0800143	1.160751
fumodurg		.377305	.1844792	2.05	0.041	.0157324	.7388776
cmorapai		-.536432	.348061	-1.54	0.123	-1.218619	.1457551
_Icbensgr_1		.1226829	.1473323	0.83	0.405	-.1660831	.4114488

**MODELO COM SEXO**

Iteration 0: log likelihood = -650.80357  
 Iteration 1: log likelihood = -649.02923  
 Iteration 2: log likelihood = -649.02444  
 Iteration 3: log likelihood = -649.02444  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233

Log likelihood = -649.02444  
 Wald chi2(8) = 840.18  
 Prob > chi2 = 0.0000

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-2.824431	.2321839	-12.16	0.000	-3.279503	-2.369359
int2		-3.061004	.2656242	-11.52	0.000	-3.581618	-2.54039
int3		-2.104609	.1781062	-11.82	0.000	-2.45369	-1.755527
int4		-1.839305	.1686626	-10.91	0.000	-2.169877	-1.508732
int5		-1.117587	.1378903	-8.10	0.000	-1.387847	-.8473269
int6		-1.213889	.1627075	-7.46	0.000	-1.53279	-.8949879
int7		-2.641835	.3398509	-7.77	0.000	-3.30793	-1.975739
sexo		-.0773759	.1314102	-0.59	0.556	-.3349351	.1801833



**MODELO COM TIPO DE PARTO**

Iteration 0: log likelihood = -650.50796  
 Iteration 1: log likelihood = -648.68255  
 Iteration 2: log likelihood = -648.67789  
 Iteration 3: log likelihood = -648.67789  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(9) = 839.67  
 Prob > chi2 = 0.0000

Log likelihood = -648.67789

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-2.882017	.2315596	-12.45	0.000	-3.335866	-2.428169
int2		-3.117137	.2649724	-11.76	0.000	-3.636473	-2.5978
int3		-2.15918	.1768077	-12.21	0.000	-2.505717	-1.812643
int4		-1.893796	.1668586	-11.35	0.000	-2.220833	-1.566759
int5		-1.172866	.1355284	-8.65	0.000	-1.438497	-.9072353
int6		-1.270116	.1604912	-7.91	0.000	-1.584673	-.9555587
int7		-2.699217	.3387752	-7.97	0.000	-3.363204	-2.03523
_Iparto_1		.1591283	.1763449	0.90	0.367	-.1865013	.5047578
_Iparto_2		-.0369993	.1564017	-0.24	0.813	-.343541	.2695424

**MODELO COM RAÇA**

Iteration 0: log likelihood = -650.86807  
 Iteration 1: log likelihood = -649.10165  
 Iteration 2: log likelihood = -649.09683  
 Iteration 3: log likelihood = -649.09683  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(8) = 840.30  
 Prob > chi2 = 0.0000

Log likelihood = -649.09683

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-2.838086	.2296417	-12.36	0.000	-3.288176	-2.387997
int2		-3.07484	.2633243	-11.68	0.000	-3.590946	-2.558734
int3		-2.118371	.1747132	-12.12	0.000	-2.460803	-1.775939
int4		-1.853679	.1646693	-11.26	0.000	-2.176424	-1.530933
int5		-1.131975	.1329562	-8.51	0.000	-1.392565	-.8713856
int6		-1.228071	.158721	-7.74	0.000	-1.539159	-.9169838
int7		-2.655831	.3380681	-7.86	0.000	-3.318432	-1.99323
racag		-.0604956	.1348425	-0.45	0.654	-.324782	.2037908

**MODELO COM IDADE DA MÃE**

Iteration 0: log likelihood = -649.83593  
 Iteration 1: log likelihood = -648.0381  
 Iteration 2: log likelihood = -648.03331  
 Iteration 3: log likelihood = -648.03331  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233

Wald chi2(9) = 838.45  
 Prob > chi2 = 0.0000

Log likelihood = -648.03331

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-2.830393	.227643	-12.43	0.000	-3.276566	-2.384221
int2		-3.066886	.2616331	-11.72	0.000	-3.579678	-2.554095
int3		-2.110756	.1718211	-12.28	0.000	-2.447519	-1.773993
int4		-1.845473	.1616488	-11.42	0.000	-2.162299	-1.528647
int5		-1.121192	.1300896	-8.62	0.000	-1.376163	-.8662214
int6		-1.213786	.1564348	-7.76	0.000	-1.520393	-.9071796
int7		-2.63828	.3367457	-7.83	0.000	-3.298289	-1.97827
_Icidadem~1		.0044791	.1632901	0.03	0.978	-.3155637	.3245218
_Icidadem~2		-.3330578	.2322458	-1.43	0.192	-.7882513	.1221357

**MODELO COM ESCOLARIDADE DA MÃE**

Iteration 0: log likelihood = -648.19067  
 Iteration 1: log likelihood = -646.33832  
 Iteration 2: log likelihood = -646.33353  
 Iteration 3: log likelihood = -646.33353  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(9) = 835.06  
 Prob > chi2 = 0.0000

Log likelihood = -646.33353

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-3.629963	.4368437	-8.31	0.000	-4.486161	-2.773765
int2		-3.865932	.4553484	-8.49	0.000	-4.758399	-2.973466
int3		-2.908772	.4100108	-7.09	0.000	-3.712379	-2.105166
int4		-2.641989	.4051948	-6.52	0.000	-3.436156	-1.847822
int5		-1.913283	.3915391	-4.89	0.000	-2.680686	-1.14588
int6		-2.003156	.3988723	-5.02	0.000	-2.784932	-1.221381
int7		-3.428865	.4972847	-6.90	0.000	-4.403525	-2.454205
_Icescolma~1		.826654	.3914428	2.11	<b>0.035</b>	.0594403	1.593868
_Icescolma~2		.7725684	.3887042	1.99	<b>0.047</b>	.0107222	1.534415

A variável ESCOLARIDADE DA MÃE deve ser selecionada.

**MODELO COM CONDIÇÃO MARITAL**

Iteration 0: log likelihood = -650.54102  
 Iteration 1: log likelihood = -648.7569  
 Iteration 2: log likelihood = -648.75211  
 Iteration 3: log likelihood = -648.75211  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(8) = 839.77  
 Prob > chi2 = 0.0000

Log likelihood = -648.75211

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-2.89625	.2268343	-12.77	0.000	-3.340837	-2.451663
int2		-3.132384	.2609403	-12.00	0.000	-3.643818	-2.620951
int3		-2.175088	.1706925	-12.74	0.000	-2.509639	-1.840537
int4		-1.909944	.1604375	-11.90	0.000	-2.224396	-1.595492
int5		-1.187191	.1276225	-9.30	0.000	-1.437327	-.937056
int6		-1.282941	.1533757	-8.36	0.000	-1.583552	-.98233
int7		-2.711691	.3350854	-8.09	0.000	-3.368446	-2.054936
_Imaritalg_2		.1500082	.1567015	0.96	0.338	-.1571211	.4571374

**MODELO COM HÁBITO DE FUMAR DA MÃE**

Iteration 0: log likelihood = -648.81241  
 Iteration 1: log likelihood = -646.91456  
 Iteration 2: log likelihood = -646.90998  
 Iteration 3: log likelihood = -646.90998  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(8) = 837.03  
 Prob > chi2 = 0.0000

Log likelihood = -646.90998

	z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1		-2.930249	.2264764	-12.94	0.000	-3.374135	-2.486364
int2		-3.163548	.2604661	-12.15	0.000	-3.674052	-2.653044
int3		-2.205242	.170009	-12.97	0.000	-2.538454	-1.872031
int4		-1.936499	.1593955	-12.15	0.000	-2.248908	-1.62409
int5		-1.207761	.1257502	-9.60	0.000	-1.454227	-.9612955
int6		-1.302138	.1518614	-8.57	0.000	-1.599781	-1.004495
int7		-2.734335	.3345565	-8.17	0.000	-3.390054	-2.078616
fumodurg		.3895246	.174503	2.23	<b>0.026</b>	.047505	.7315442

A variável HÁBITO DE FUMAR deve ser selecionada.

**MODELO COM MORAR O PAI DA CRIANÇA COM A MÃE**

Iteration 0: log likelihood = -650.92441  
 Iteration 1: log likelihood = -649.15615  
 Iteration 2: log likelihood = -649.15134  
 Iteration 3: log likelihood = -649.15134  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(8) = 840.39  
 Prob > chi2 = 0.0000

Log likelihood = -649.15134

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1	-2.871556	.2259205	-12.71	0.000	-3.314352	-2.42876
int2	-3.107913	.2601467	-11.95	0.000	-3.617791	-2.598035
int3	-2.151516	.169632	-12.68	0.000	-2.483989	-1.819044
int4	-1.886823	.1594855	-11.83	0.000	-2.199409	-1.574237
int5	-1.164769	.1265753	-9.20	0.000	-1.412852	-.9166859
int6	-1.261534	.1526917	-8.26	0.000	-1.560804	-.962264
int7	-2.690672	.3348266	-8.04	0.000	-3.34692	-2.034424
cmorapai	.0519447	.1692649	0.31	0.759	-.2798084	.3836978

**MODELO COM POSSE DE BENS**

Iteration 0: log likelihood = -650.43954  
 Iteration 1: log likelihood = -648.64548  
 Iteration 2: log likelihood = -648.64071  
 Iteration 3: log likelihood = -648.64071  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(8) = 839.58  
 Prob > chi2 = 0.0000

Log likelihood = -648.64071

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1	-2.920628	.2309511	-12.65	0.000	-3.373284	-2.467973
int2	-3.15668	.2644894	-11.93	0.000	-3.675069	-2.63829
int3	-2.199519	.1760296	-12.50	0.000	-2.544531	-1.854507
int4	-1.933881	.1659197	-11.66	0.000	-2.259077	-1.608684
int5	-1.211342	.1344404	-9.01	0.000	-1.47484	-.9478435
int6	-1.307159	.1590148	-8.22	0.000	-1.618822	-.9954953
int7	-2.735834	.3376775	-8.10	0.000	-3.39767	-2.073998
_Icbensgr_1	.1412374	.1332185	1.06	0.289	-.1198661	.4023408

**MODELO COM VARIÁVEIS SELECIONADAS - ESCOLARIDADE DA MÃE E HÁBITO DE FUMAR**

Iteration 0: log likelihood = -646.38896  
 Iteration 1: log likelihood = -644.41766  
 Iteration 2: log likelihood = -644.4131  
 Iteration 3: log likelihood = -644.4131  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(10) = 832.25  
 Prob > chi2 = 0.0000

Log likelihood = -644.4131

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1	-3.640414	.4369749	-8.33	0.000	-4.496869	-2.783959
int2	-3.873587	.4554372	-8.51	0.000	-4.766228	-2.980947
int3	-2.914927	.4100913	-7.11	0.000	-3.718691	-2.111163
int4	-2.644742	.4052386	-6.53	0.000	-3.438995	-1.850489
int5	-1.910092	.3915209	-4.88	0.000	-2.677459	-1.142725
int6	-1.998083	.3988237	-5.01	0.000	-2.779763	-1.216403
int7	-3.426818	.4972639	-6.89	0.000	-4.401437	-2.452199
_Icescolma~1	.7823366	.3922625	1.99	0.086	.0135163	1.551157
_Icescolma~2	.7113256	.3902607	1.82	0.098	-.0535714	1.476223
fumodurg	.3572137	.1754492	2.04	0.042	.0133395	.7010879

**MODELO COM HÁBITO DE FUMAR**

Iteration 0: log likelihood = -648.81241  
 Iteration 1: log likelihood = -646.91456  
 Iteration 2: log likelihood = -646.90998  
 Iteration 3: log likelihood = -646.90998  
 Complementary log-log regression

Number of obs = 1879  
 Zero outcomes = 1646  
 Nonzero outcomes = 233  
 Wald chi2(8) = 837.03  
 Prob > chi2 = 0.0000

Log likelihood = -646.90998

z	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
int1	-2.930249	.2264764	-12.94	0.000	-3.374135	-2.486364
int2	-3.163548	.2604661	-12.15	0.000	-3.674052	-2.653044
int3	-2.205242	.170009	-12.97	0.000	-2.538454	-1.872031
int4	-1.936499	.1593955	-12.15	0.000	-2.248908	-1.62409
int5	-1.207761	.1257502	-9.60	0.000	-1.454227	-.9612955
int6	-1.302138	.1518614	-8.57	0.000	-1.599781	-1.004495
int7	-2.734335	.3345565	-8.17	0.000	-3.390054	-2.078616
fumodurg	.3895246	.174503	2.23	0.026	.047505	.7315442

O TESTE DA RAZÃO DE VEROSSIMILHANÇAS, A PARTIR DOS DOIS MODELOS ANTERIORES (UM COM HÁBITO DE FUMAR E ESCOLARIDADE DA MÃE E O OUTRO SOMENTE COM HÁBITO DE FUMAR), INDICA A SAÍDA DA VARIÁVEL ESCOLARIDADE DA MÃE.

chi2(1) = 4,99  
 Prob > chi2 = 0,0823

DESSA FORMA, O MODELO FINAL FOI O QUE CONTÉM SOMENTE HÁBITO DE FUMAR DA MÃE.

**Anexo 8 – Programa no STATA para o modelo complemento log-log  
e alterações no banco de dados**

### Programa para preparar os dados para rodar o modelo complemento log-log.

```

*
* complemento log log
*

* abrir arquivo
use "C:\bancos\banco.dta", clear

* criar os intervalos a partir da informação registrada
gen interv = recode(peito, 15, 30, 60, 90, 180, 270, 360)

* criando-se os intervalos de 1 a 7 para replicar
gen interval = 1 if interv == 15
replace interval = 2 if interv == 30
replace interval = 3 if interv == 60
replace interval = 4 if interv == 90
replace interval = 5 if interv == 180
replace interval = 6 if interv == 270
replace interval = 7 if interv == 360

* utiliza-se a variável interval para replicar os dados,
* segundo o número de intervalos
expand interval

* criar a variável periodo que indica o número total de intervalos
* para cada indivíduo
sort n5cri
by n5cri: gen periodo = _n

* criar a variável indicadora c. É preciso se considerar o indivíduo
* (n5cri) e o status (stapei).
gen c = 0 if stapei==0
replace c = 0 if n5cri[_n]==n5cri[_n+1] & stapei==1
replace c = 1 if n5cri[_n]~=n5cri[_n+1] & stapei==1

* criar o vetor "tau" (int1, ..., int7) do modelo
gen int1=1 if periodo==1
replace int1=0 if periodo~=1
gen int2=1 if periodo==2
replace int2=0 if periodo~=2
gen int3=1 if periodo==3
replace int3=0 if periodo~=3
gen int4=1 if periodo==4
replace int4=0 if periodo~=4
gen int5=1 if periodo==5
replace int5=0 if periodo~=5
gen int6=1 if periodo==6
replace int6=0 if periodo~=6
gen int7=1 if periodo==7
replace int7=0 if periodo~=7

* o STATA possui o comando "cloglog" para essa regressão
cloglog c int1 int2 int3 int4 int5 int6 int7, noconstant
* ou, por meio do comando glm (modelos lineares generalizados),
glm c int1 int2 int3 int4 int5 int6 int7, family(binomial)
    link(cloglog) noconstant

```

O Exemplo a seguir apresenta um fragmento do banco de dados após executar o programa no software STATA.

Exemplo. banco de dados após o programa

id	interval	periodo	c	int1	int2	int3	int4	int5	int6
55	1	1	1	1	0	0	0	0	0
64	2	1	0	1	0	0	0	0	0
64	2	2	1	0	1	0	0	0	0
21	3	1	0	1	0	0	0	0	0
21	3	2	0	0	1	0	0	0	0
21	3	3	1	0	0	1	0	0	0
2	4	1	0	1	0	0	0	0	0
2	4	2	0	0	1	0	0	0	0
2	4	3	0	0	0	1	0	0	0
2	4	4	0	0	0	0	1	0	0
7	5	1	0	1	0	0	0	0	0
7	5	2	0	0	1	0	0	0	0
7	5	3	0	0	0	1	0	0	0
7	5	4	0	0	0	0	1	0	0
7	5	5	1	0	0	0	0	1	0
1	6	1	0	0	0	0	0	0	0
1	6	2	0	0	1	0	0	0	0
1	6	3	0	0	0	1	0	0	0
1	6	4	0	0	0	0	1	0	0
1	6	5	0	0	0	0	0	1	0
1	6	6	0	0	0	0	0	0	1

A criança 55 deixou o leite de peito (c=1) no intervalo 1.

A criança 64 não deixou o leite de peito (c=0) no intervalo 1 (periodo=1) mas deixou o leite de peito (c=1) no intervalo 2 (periodo=2).

A criança 21 deixou o leite de peito (c=1) no intervalo 3 e não deixou (c=0) até o intervalo 2 (periodo=2).

A criança 2 deixou o leite de peito (c=1) no intervalo 4 e não deixou (c=0) até o intervalo 3 (periodo=3).

A criança 7 deixou o leite de peito (c=1) no intervalo 5 e não deixou (c=0) até o intervalo 4 (periodo=4).

A criança 1 não deixou o leite de peito (c=0) até o intervalo 6, quando houve censura por terminar o seguimento.

**Anexo 9 – Parecer do Comitê de Ética em Pesquisa da FSP - COEP**