UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE MEDICINA SOCIAL

ADRIANA DE FÁTIMA LOURENÇON

**Modelos de Análise de Sobrevida aplicados ao Estudo do
Comportamento de Retorno do Doador de Sangue**

Ribeirão Preto
2007

ADRIANA DE FÁTIMA LOURENÇON

**Modelos de Análise de Sobrevida aplicados ao Estudo do
Comportamento de Retorno do Doador de Sangue**

Dissertação de Mestrado apresentada à Pós
Graduação da Faculdade de Medicina de
Ribeirão Preto da Universidade de São Paulo,
para obtenção do título de Mestre em Ciências
Médicas, junto ao Departamento de Medicina
Social.

Área de Concentração: Saúde na Comunidade

Orientador: Prof. Dr. Edson Zangiacomi Martinez

Ribeirão Preto
2007

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTA TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

LOURENÇON, ADRIANA DE FÁTIMA

Modelos de Análise de Sobrevivência aplicados ao estudo do comportamento de retorno do doador de sangue. / Adriana de Fátima Lourençon; Orientador: Prof. Dr. Edson Zangiacomi Martinez. --
Ribeirão Preto, 2007.

154 f.: il.; 30 cm

Dissertação de Mestrado apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para a obtenção do título de Mestre, Área de Concentração em Saúde da Comunidade.

1. Análise de Sobrevivência, 2. Modelos de Longa-Duração, 3. Doador de Sangue, 4. Comportamento de Retorno.

FOLHA DE APROVAÇÃO

Adriana de Fátima Lourençon

Modelos de Análise de Sobrevivência aplicados ao estudo do comportamento de retorno do doador de sangue.

Dissertação de Mestrado apresentada à Pós Graduação da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências Médicas, junto ao Departamento de Medicina Social.

Aprovado em: ____/____/____

Banca Examinadora

Prof(a). Dr(a).: _____

Instituição: _____

Prof(a). Dr(a).: _____

Instituição: _____

Prof(a). Dr(a).: _____

Instituição: _____

⇒ Trabalho apresentando em congresso:

Estudo do Comportamento de Retorno do Doador de Sangue Voluntário utilizando Modelos de Sobrevivência de Longa-Duração.

Adriana de Fátima Lourençon, Oranice Ferreira, Josmar Mazucheli, Edson Zangiacomi Martinez.

10º Escola de Modelos de Regressão, Salvador, BA, Fevereiro/2007.

⇒ Artigo publicado:

Modelos de Sobrevivência de Longa-Duração aplicados ao estudo do comportamento de retorno do doador de sangue voluntário.

Adriana de Fátima Lourençon, Oranice Ferreira, Josmar Mazucheli, Edson Zangiacomi Martinez.

Revista Brasileira de Matemática e Estatística, São Paulo, v. 25, n. 1, p. 137-154, 2007.

⇒ Artigo em fase de submissão:

Taxas de retorno de doadores de sangue de primeira vez do Centro Regional de Hemoterapia de Ribeirão Preto.

Adriana de Fátima Lourençon, Oranice Ferreira, Amilton Gomes de Brito, Edson Zangiacomi Martinez.

Revista Brasileira de Hematologia e Hemoterapia, 2007.

Aos meus pais,
Valdir Lourençon e
Aparecida Caetano de Oliveira Lourençon,
Pelo incentivo e pelo exemplo de força e garra
os quais me impulsionaram para realização dessa conquista.

AGRADECIMENTOS

Á DEUS, força maior de todo ser humano.

Aos meus pais, *Aparecida Caetano de Oliveira Lourençon* e *Valdir Lourençon*, pelo amor e pelo ensinamento de que o mais importante é o que somos e não o que temos, e as minhas irmãs, *Andréia de Cássia Lourençon* e *Alessandra de Paula Lourençon*, meus grandes exemplos de vida.

Aos professores Prof. Dr. Afonso Dinis Costa Passos e Prof. Dr. Francisco Louzada-Neto pelas sugestões e em especial ao meu orientador Prof. Dr. Edson Zangiacomi Martinez que esteve presente ao longo deste período de aprendizado.

À todos os funcionários do Hemocentro de Ribeirão Preto por toda cooperação demonstrada, principalmente ao Amilton Gomes de Brito (Centro de Processamento de Dados) e a minha grande amiga Oranice Ferreira (Seção de Enfermagem), pela sua paciência e auxílio em esclarecer todas as minhas dúvidas. Agradeço ainda o apoio recebido do Prof. Dr. Dimas Tadeu Covas, diretor científico do Hemocentro.

À todos os meus amigos que me entenderam, me apoiaram e me deram muitas alegrias durante esses anos de mestrado. Distantes ou próximos, sempre tiveram uma palavra amiga e certa em todos os momentos. Eliza Omai, Silvia Regina Lamas, Clécia Aparecida Garcia, vocês são inesquecíveis.

À vocês, Denise Nascimento Mesquita e Ailton Massaiti Watanabe, peças contribuintes e indispensáveis neste trabalho, o meu muito obrigado.

A você Ailton Massaiti Watanabe, que lado a lado, acompanhou este trabalho e com quem agora tenha a imensa alegria de dividir mais essa vitória.

Às secretárias do Departamento de Medicina Social, Carolina Cecília Bruno Batista, Regina Helena Alcântara e em especial a Mônica Elisabet Knack por todo trabalho realizado com tanta dedicação. Também agradeço a Solange Pedersoli, pela atenção e disponibilidade.

A Fundação de Apoio ao Ensino, Pesquisa e Assistência do Hospital das Clínicas da FMRP/USP (FAEPA - HCFMRP) pelo apoio financeiro.

Um forte abraço a todos!

*“Pobre de quem teve medo de correr riscos,
porque talvez não se decepcionou nunca,
nem teve desilusões,
nem sofreu como aqueles que têm um sonho a seguir, ...,
Então, esta será tua herança.
A certeza de que desperdiçaste a vida.”*

(Autor Desconhecido)

RESUMO

LOURENÇON, A.F. **Modelos de Análise de Sobrevida aplicados ao Estudo do Comportamento de Retorno do Doador de Sangue**. 2007. 154 f. Dissertação de Mestrado – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2007.

Notícias de escassez no mundo inteiro, dada a crescente demanda e o rigor na triagem clínica, levaram a necessidade de investigar métodos que mensurem o comportamento de retorno do doador de sangue, sobretudo o indivíduo que manifesta a intenção voluntária em doar. Curvas de Sobrevida entre outros métodos estatísticos são amplamente estudados na literatura com o intuito de obter uma estimativa da chance de um doador vir a realizar uma subsequente doação, associado ao seu perfil. O objetivo do presente estudo é identificar modelos estatísticos capazes de descrever esse comportamento utilizando os registros do Centro Regional de Hemoterapia de Ribeirão Preto. A construção de modelos de longa-duração, por exemplo, pode ser um meio de evidenciar possíveis subgrupos mais propensos a retornar, além de estimar a proporção de doadores que jamais retornarão. Entre os resultados, observamos que apenas 40% dos doadores voluntários retornaram após um ano decorrido da primeira doação, e 20% destes jamais retornarão. O ajuste do modelo longa-duração possibilitou ainda indicar alguns subgrupos de doadores prováveis e improváveis de retornar, porém tais resultados reforçam as evidências de que a motivação intrínseca é o que leva o indivíduo a retornar.

Palavras-chave: Análise de Sobrevida, Modelos de Longa-Duração, Doador de Sangue, Comportamento de Retorno.

ABSTRACT

LOURENÇON, A.F. **Survival Analysis Models applied to the Study of Blood Donor Return Behavior**. 2007. 154 p. Dissertation (Master Degree). Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2007.

Reports of worldwide shortages due the increased demand and rigor of clinical screening have led to the necessity to investigate methods that measure blood donor return behavior, mainly regarding individuals who manifest the voluntary intention to donate. Survival curves, among others statistical methods, have been extensively studied in the literature in order to estimate the likelihood of a donor to make another donation, associated with his profile. The aim of the present study was to identify statistical models describing this behavior using information from the Regional Hemotherapy Center of Ribeirão Preto. The construction of long-term survival model can be a useful instrument for determining the groups more likely to donate, as well as the proportion of donors who will never return. The results obtained revealed that only 40% of the volunteer primary donors return for a new donation one year after the first, with the estimate that 20% will never return. The construction of long-term survival model still facilitated to indicate some groups likely and unlikely donors to donate, even so such results reinforce the evidences that the intrinsic motivation is what prompts a donor to return.

Keywords: Survival Analysis, Long-Term Survival Models, Blood Donor, Return Behavior.

Lista de Tabelas

CAPÍTULO 4

Resultados

Tabela 4.3.1.	Descrição das variáveis sócio-demográficas dos doadores voluntários do Centro Regional de Hemoterapia de Ribeirão Preto.	74
Tabela 4.3.2.	Distribuição do tipo sanguíneo e fator Rh dos doadores voluntários do Centro Regional de Hemoterapia de Ribeirão Preto e do esperado encontrar na população.	76
Tabela 4.3.3.	Distribuição dos doadores voluntários que retornaram para sua segunda doação de acordo com o tipo de doação.	77
Tabela 4.3.4.	Distribuição dos doadores voluntários que retornaram para sua segunda doação de acordo com a condição de triagem.	77
Tabela 4.3.5.	Distribuição dos doadores voluntários que retornaram para sua segunda doação de acordo com a unidade.	78
Tabela 4.3.6.	Descrição do tempo entre a primeira e a segunda doação do doador voluntário de acordo com o número de retornos realizados ao longo do histórico de doação.	79
Tabela 4.3.7.	Distribuição do tempo entre a primeira e a segunda doação ou até o fim do seguimento o doador voluntário.	80
Tabela 4.4.1.	Sumário das estimativas de Kaplan-Meier para cada uma das variáveis sócio-demográficas dos doadores voluntários.	86
Tabela 4.5.1.	Parametrização das distribuições 1)Weibull, 2) Log-Logística, 3)Log-Normal e 4)Gama Generalizada Estendida.	92
Tabela 4.5.2.	Estimativas dos parâmetros dos modelos ajustados assumindo as distribuições de 1)Weibull, 2) Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida juntamente como os valores obtidos para a Log-Verossimilhança e critério BIC.	93
Tabela 4.5.3.	Estimativas dos Parâmetros dos Modelos de Longa-Duração ajustados assumindo as distribuições de 1)Weibull, 2)Log-Logística, 3)Log-Normal e 4)Gama Generalizada Estendida juntamente como os	97

valores obtidos para o Log-Verossimilhança e critério BIC.

Tabela 4.6.1.	Estimativas dos parâmetros do modelo de regressão de longa-duração assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação para cada covariável relacionada com o perfil do doador voluntário – Ajuste I.	104
Tabela 4.6.2	Proporção esperada de doadores voluntários improváveis de retornar de acordo com as covariáveis relacionadas com o seu perfil a partir do Ajuste I.	107
Tabela 4.6.3.	Valores do critério BIC para os modelos ajustados assumindo uma distribuição 1)Weibull, 2)Log-Logística, 3)Log-Normal e 4)GGE para o tempo entre a primeira e a segunda doação.	108
Tabela 4.6.4.	Estimativas dos parâmetros do modelo de regressão de longa-duração assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação para cada covariável relacionada com o perfil do doador voluntário – Ajuste II.	109
Tabela 4.6.5.	Valores da maximização do Log-Verossimilhança e do critério BIC dos modelos assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação.	111

Lista de Figuras

CAPÍTULO 3

Sujeito e Métodos

- Figura 3.4.1.** Casos particulares da distribuição Gama Generalizada Estendida (GGE). 50
- Figura 3.4.2.** Função de sobrevivência de um modelo paramétrico assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação do doador atribuindo diversos valores para os seus parâmetros. 51
- Figura 3.4.3.** Função de sobrevivência de um modelo paramétrico de longa-duração assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação do doador atribuindo diversos valores para os seus parâmetros. 61

CAPÍTULO 4

Resultados

- Figura 4.1.1.** Número de doações mensais realizadas no período de 1996 a 2005 obtidas na região de abrangência de Ribeirão Preto. 66
- Figura 4.1.2.** Número de doações mensais realizadas no período de 1996 a 2005 obtidas na região de abrangência de Ribeirão Preto de acordo com o tipo de doação: autóloga (1º), específica (2º), vinculada e voluntária (3º). 67
- Figura 4.1.3.** Número de doadores que realizaram sua primeira doação mensalmente no período de 1996 a 2005 obtidas na região de abrangência de Ribeirão Preto. 68
- Figura 4.1.4.** Número de doadores mensais registrados no período de 1996 a 2005 da região de abrangência de Ribeirão Preto de acordo com o tipo de doação: autóloga (1º), específica (2º), vinculada e voluntária (3º). 70
- Figura 4.2.1.** Número de retornos realizados no período de 1996 a 2005 de acordo com o tempo de acompanhamento dos doadores voluntários. 72
- Figura 4.3.1.** Porcentagem de doadores voluntários de acordo com o número total de retornos realizados e o tempo entre a primeira e a segunda doação. 79

Figura 4.3.2.	Porcentagem acumulada do tempo entre a primeira e a segunda doação do doador voluntário ou até o fim do seguimento o doador voluntário.	80
Figura 4.4.1.	Função de sobrevivência representando a propensão à não retornar dos doadores voluntários.	82
Figura 4.4.2.	Função risco instantâneo à retornar. Os picos mais expressivos de retorno correspondem aproximadamente a 4, 12 e 24 meses após a primeira doação.	82
Figura 4.4.3.	Função de sobrevivência representando a propensão à não retornar dos doadores voluntários segundo unidade de coleta.	84
Figura 4.4.4.	Função risco instantâneo à retornar segundo unidade de coleta. Os picos mais expressivos de retorno correspondem aproximadamente a 4, 12 e 24 meses após a primeira doação.	84
Figura 4.4.5.	Função de sobrevivência representando a propensão à não retornar dos doadores voluntários segundo as variáveis sócio-demográficas. (1- Sexo, 2- Cor da Pele, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh).	87
Figura 4.4.6.	Função risco instantâneo à retornar dos doadores voluntários segundo as variáveis sócio-demográficas. (1- Sexo, 2- Cor, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh).	89
Figura 4.4.7.	Porcentagem de retorno de doadores voluntários em cada categoria das variáveis sócio-demográficas. A notação de cada categoria corresponde as suas iniciais. A linha de referencia corresponde ao número de doadores que retornaram para uma nova doação independente de qualquer variável sócio-demográfica.	90
Figura 4.5.1.	Curvas de sobrevivência estimadas pelos modelos paramétricos assumindo as distribuições: 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida seguida da curva estimada por Kaplan-Meier.	94
Figura 4.5.2	Gráficos das funções de sobrevivência estimadas por Kaplan-Meier <i>versus</i> as funções de sobrevivência estimadas pelos modelos assumindo uma distribuição 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida.	96
Figura 4.5.3.	Curvas de sobrevivência estimadas pelos modelos de longa-duração assumindo as distribuições de 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida, seguida da curva estimada por Kaplan-Meier.	98

Figura 4.5.4.	Gráficos das funções de sobrevivência estimadas por Kaplan-Meier <i>versus</i> as funções de sobrevivência estimadas pelos modelos de longa-duração assumindo uma distribuição 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida.	100
Figura 4.6.1.	Logaritmo da função risco instantâneo à retornar acumulado dos doadores voluntários segundo a variável unidade de coleta.	102
Figura 4.6.2.	Logaritmo da função risco instantâneo à retornar acumulado dos doadores voluntários segundo as variáveis sócio-demográficas. (1- Sexo, 2-Cor, 3-Faixa Etária, 4-Estado Civil, 5- Tipo Sanguíneo, 6-Fator Rh).	103
Figura 4.6.3.	Função de sobrevivência representando a propensão à não retornar dos doadores voluntários segundo as variáveis sócio-demográficas e os respectivos ajustes pelo modelo de regressão de longa duração com distribuição GGE. (1- Sexo, 2- Cor, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh) – Ajuste I.	105
Figura 4.6.4.	Função de sobrevivência representando a propensão à não retornar dos doadores de sangue voluntários de primeira vez segundo unidade de coleta e respectivo ajuste pelo modelo de longa-duração com distribuição GGE.	106
Figura 4.6.5.	Resultados da área sob a curva ROC obtidos com os ajustes dos modelos assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação de acordo com o número de covariáveis.	114

SUMÁRIO

CAPÍTULO 1.....	17
Introdução	
1.1 As Políticas de Sangue no Brasil	25
1.2 O Centro Regional de Hemoterapia de Ribeirão Preto	28
1.3 A doação de sangue: conceitos e definições	29
 CAPÍTULO 2.....	 31
Objetivos	
2.1 Geral	32
2.2 Específicos	32
 CAPÍTULO 3.....	 33
Sujeito e Métodos	
3.1 Casuística	34
3.2 Critérios de Inclusão/Exclusão	34
3.3 Variáveis	36
3.4 Metodologia	38
3.4.1 Análise de Sobrevivência	39
3.4.1.1 Estimador não paramétrico de Kaplan-Meier	42
3.4.1.2 Modelos Paramétricos em Análise de Sobrevivência	43
3.4.1.3 Modelos de Regressão Paramétricos em Análise de Sobrevivência	52
3.4.1.4 Estimação de Modelos de Regressão Paramétricos	53
3.4.1.5 Modelos de Regressão Paramétricos de Longa-Duração em Análise de Sobrevivência	55
3.4.1.6 Estimação de Modelos de Regressão de Longa-Duração Paramétricos	62
3.4.1.7 Medidas de avaliação de Modelos em Análise de Sobrevivência .	63

CAPÍTULO 4.....	65
Resultados	
4.1 Descrição do Histórico de Doação	66
4.2 Descrição do Histórico de Doação dos Doadores Voluntários	71
4.3 Descrição do Número de Doadores Voluntários de acordo com as variáveis sócio-demográficas	73
4.4 Descrição do tempo entre a primeira e a segunda doação dos doadores voluntários	81
4.5 Aplicação de Modelos de Análise de Sobrevida utilizando o tempo entre a primeira e a segunda doação dos doadores voluntários	91
4.6 Aplicação de Modelos de Análise de Sobrevida utilizando o tempo entre a primeira e a segunda doação dos doadores voluntários e covariáveis ..	101
 CAPÍTULO 5.....	 115
Discussão	
 CAPÍTULO 6.....	 123
Conclusão	
 CAPÍTULO 7.....	 126
Perspectivas Futuras	
 REFERENCIAS BIBLIOGRÁFICAS.....	 128
 ANEXOS.....	 135

CAPÍTULO I

Introdução

A procura de sangue e hemoderivados, garantindo seu fornecimento de maneira segura e sustentável, é um dos maiores desafios enfrentados pelos serviços de hemoterapia.

Com o aumento da população idosa, do acesso da população aos cuidados médicos e da complexidade dos seus procedimentos, a preocupação com a escassez de sangue se torna evidente (BRASIL, 2004b), tornando-se necessário a realização de programas que visem o fornecimento seguro de sangue por meio da captação e fidelização de doadores voluntários provenientes de populações de baixo risco.

Ludwig e Rodrigues (2005) enfatizam que a busca do doador voluntário e habitual se deve principalmente a razões econômicas e de segurança, dado que doadores testados e re-testados implicam em bolsas de sangue com maior margem de segurança para o receptor e um número menor de exames sorológicos desprezados.

Notícias divulgadas pelo Ministério da Saúde (BRASIL, S.d.a) refletem essa preocupação em vários Hemocentros do Brasil, onde se observa a queda do número de doadores e de doações principalmente em determinadas épocas do ano (período de férias), o que coincide com um aumento da demanda por bolsas de sangue devido o maior tráfego nas rodovias e um conseqüente aumento no número de acidentes com vítimas que necessitam de transfusões. Além disso, o crescimento da violência também é um fator apontando para esse aumento na demanda por bolsas de sangue, principalmente nas grandes cidades.

Além do fator sazonal na demanda por bolsas de sangue, Simon (2003) aponta outro agravante à diminuição do número de candidatos à doação e conseqüentemente na oferta de bolsas de sangue: o rigor na triagem clínica dos doadores e a introdução de novas causas de inaptidão. Um exemplo disso é a polêmica gerada em torno da doação de homens que tiveram relações sexuais com outros homens, no qual a ANVISA considera como um público inapto temporariamente a doar (BRASIL, 2004a, 2006).

Cita ainda, que a incompatibilidade sanguínea é um fator que implica na disponibilidade de sangue e hemocomponentes, de tal maneira que a oferta de alguns grupos pode ser mais difícil de sustentar, pois são muito transfundidos, como é o caso do sangue tipo O. O indivíduo portador desse tipo sanguíneo é considerado doador universal e portanto pode doar para todos os demais tipos sanguíneos.

Em números, estudos vêm evidenciando que essa parcela de predispostos a doar é ainda reduzida durante o processo de doação de sangue, devido à inaptidão definitiva ou temporária, sendo este último associado a um público mais propenso a não retornar conforme mostram os estudos de Pilliavin (1987) e Halperin, Baetens e Newman (1998).

A inaptidão definitiva ocorre quando o sangue de um determinado indivíduo é considerado inutilizado por alguma razão, entre elas o exame laboratorial positivo para doenças transmissíveis. Já a inaptidão temporária, o indivíduo é temporariamente impossibilitado de doar por não passar por alguma etapa no processo de triagem.

O processo de triagem é dividido em várias etapas: após a pré-triagem, onde são verificadas a pressão arterial, pulso, temperatura corporal, peso, altura e hemoglobina (diagnóstico ou não de anemia), o doador é entrevistado, em condições de privacidade, por um profissional da área de saúde¹. Posteriormente é realizada a coleta de sangue e servido um lanche.

Outro aspecto a ser considerado neste processo, que antecede ao ciclo de doação, é o entendimento das razões que levam os indivíduos a doar sangue ou não. Neste sentido, Montoya et al. (1996) conduziram um estudo entre doadores e não doadores de Granada, Espanha, onde destacaram que o medo e a falta de informação são os fatores mais frequentes entre os não doadores. A apatia, a desconfiança e a falta de acessibilidade aos

¹ Os critérios utilizados para aprovação ou rejeição concordam com os exigidos pelo Ministério de Saúde - Portaria 1376/93 e 121/95.

postos de coleta, também foram fatores presentes. Já entre os indivíduos que passaram pela experiência de doar, os fatores negativos tendem a ser mais amenos, sendo apontados fatores relacionados com a acessibilidade e conforto. A solidariedade, a informação e os benefícios familiares e pessoais foram as principais motivações encontradas entre eles.

O reflexo dos fatores desmotivadores também é demonstrado quando o indivíduo pretende doar pela primeira vez, tendendo a apresentar níveis de ansiedade mais elevados e medo do que doadores que já passaram pela experiência de doar, conforme visto por Borges et al. (2005) em um estudo realizado no Hemocentro de Ribeirão Preto, São Paulo. Destacam ainda que a principal motivação da doação encontrada entre os novos doadores foi o fato de ser um pedido de parentes ou amigos.

Tal motivação também foi observada por Passos e Farina (1994) num estudo realizado entre acadêmicos da Faculdade de Medicina de Ribeirão Preto que já haviam doado sangue anteriormente, além da solicitação durante o serviço militar. Destaca-se também, entre aqueles que não haviam doado, a principal razão por não fazê-lo, foi o fato de nunca ter ocorrido solicitação.

Por sua vez, Suárez et al. (2004), com o intuito de estudar as motivações dos doadores regulares de Granada, Espanha, distinguiram três pontos de vista: (a) aqueles que reconhecem a necessidade são emotivos e tendem a continuar a doar sangue, motivados pela ação altruísta, (b) aqueles que expressam uma colaboração em geral, são mais racionais e consideram a doação um caminho para ajudar as pessoas mais facilmente, manifestando a intenção de continuar a doar sempre que possível, e (c) aqueles que apresentam outras motivações menos “concretas” onde a frequência da doação é estabilizada de acordo com o acesso aos pontos de coleta.

Entre os doadores fidelizados, Moura et al. (2006) observaram em um estudo realizado no Hemocentro Regional de Crato, Ceará, as suas principais motivações na primeira

doação. Fatores altruístas, tais como solidariedade e necessidade na família foram os mais apontados.

Ainda que esses aspectos intangíveis ofereçam um entendimento das razões que levam um indivíduo a doar sangue, identificar subgrupos de doadores mais propensos ou menos propensos a retornar à doação pode ser a base para que os serviços de hemoterapia direcionem estratégias diferenciadas para reter e incentivar doadores aptos à doação.

Com este propósito, James e Matthews (1993), sugeriram os ciclos de doações como um tipo de estrutura para medir o comportamento de retorno do doador.

O ciclo de doação é definido como uma série de quatro eventos subsequentes: o indivíduo inicia seu ciclo a partir da primeira intenção a doar (primeira doação), depois, há um espaço de tempo mínimo necessário para que o seu organismo recupere, seguindo um intervalo de tempo escolhido para retornar e finaliza com uma subsequente tentativa de doação (segunda doação); simultaneamente o próximo ciclo de doação se inicia até que o indivíduo seja desqualificado de novas doações.

Assim, o ciclo de doação se preocupa com o tempo entre as sucessivas intenções de doação, sejam elas aprovadas ou um caso de inaptidão temporária do doador. Com isso, múltiplos ciclos para um mesmo indivíduo vão sendo formados e distinguidos por meio de um processo de contagem.

Neste caso, estimativas da função de sobrevivência obtidas pelo método de Kaplan-Meier, são amplamente discutidas na literatura (JAMES; MATTHEWS, 1993, 1996; OWNBY; WATANABE; NASS, 1999), com o intuito de estimar a propensão a retornar dos indivíduos que, após iniciarem o seu ciclo, retornam e dão início a um novo ciclo após um determinado período de acompanhamento. Além disso, podemos utilizar suas estimativas para descrever o efeito de uma determinada covariável relacionada com o perfil do doador, por

exemplo, o sexo, e assim verificar o comportamento de retorno ao longo do tempo em cada categoria.

Quando o interesse está na relação do tempo entre os dois primeiros ciclos de doação e as covariáveis relacionadas com o perfil do doador ou o histórico de doação, o método comumente empregado é baseado nos modelos de riscos proporcionais introduzidos por Cox (JAMES; MATTHEWS, 1996; OWNBY; WATANABE; NASS, 1999).

Outro método muito estudado é a estimação da Função Risco ou *Hazard Function*, sendo neste contexto, um indicador do comportamento de retorno do doador associado com o tamanho do intervalo (dias, semanas,...) desde a última doação (JAMES; MATTHEWS, 1993, 1996). Em outras palavras, representa a “chance” do doador retornar em um determinado período.

Muitos estudos têm evidenciado algumas características sócio-demográficas bem como seu histórico de doação como fatores preditivos na segmentação desse público em subgrupos mais ou menos propensos a retornar a doar sangue.

James e Matthews (1993), por exemplo, com o objetivo de estudar o comportamento de retorno do doador de sangue tipo O utilizando os registros do *Gulf Coast Regional Blood Center* (GCRBC) em Houston, Texas, indicaram que doadores com fator Rh negativo, tendem a retornar mais cedo do que os doadores Rh positivo. Esse fator é o mais utilizado nos serviços de hemoterapia e por isso estão sempre em falta.

James e Matthews (1996), numa segunda etapa, buscando explorar os fatores relacionados com o comportamento do doador de sangue do tipo O, observaram que o sexo, a idade, o fator Rh e o fato de ser um doador raro ou considerado temporariamente inapto, são as variáveis mais associadas com o retorno do doador. A partir da construção de um modelo de Cox observaram que determinados subgrupos de doadores tendem a retornar mais cedo, mas depende do tempo decorrido do início da doação. Assim, jovens doadores, ou indivíduos

do sexo feminino ou portadores de Rh negativo, são mais propensos a retornar para uma subsequente doação apenas durante o primeiro ano decorrido da última doação. Ser um doador raro é um fator favorável ao retorno do doador em todo o período estudado. Por outro lado, ser um doador temporariamente inapto indica um público menos propenso a retornar. Os resultados também evidenciaram a influência do histórico de doação² no retorno do indivíduo, estando associado com intervalos menores entre as subseqüentes intenções de doar.

Ownby, Watanabe e Nass (1999), ao verificar o comportamento de retorno de uma coorte de doadores de sangue de vários centros de *Retrovirus Epidemiology Donor Study* (REDS) dos Estados Unidos, também observaram que o número de doações anteriores é inversamente relacionado ao tempo entre as primeiras doações. A partir da construção de um modelo de Cox para cada ciclo de doação finalizado, concluíram que doadores portadores de Rh negativo, declarados da cor branca, com idade superior a 30 anos, nível de escolaridade alto ou histórico de transfusão, tendem a retornar mais cedo do que seus complementares, principalmente nos primeiros ciclos. A inclusão da interação nível de escolaridade e sexo revelou que doadores do sexo feminino e nível de escolaridade alto tendem a retornar mais cedo do que doadores do sexo masculino e mesmo nível de escolaridade.

Whyte (1999), com o intuito de descrever o comportamento de retorno do doador de sangue, conduziu um estudo utilizando os registros de doadores voluntários de Melbourne, Austrália, e concluiu que a proporção de doadores que retornaram em dois anos aumenta de acordo com o histórico de doação³ e com a idade.

Flegel, Besenfelder e Wagner (2000), também mostraram que o histórico de doação⁴ é a variável mais importante para predizer a probabilidade de retornar a doar entre os

² James e Matthews, nesse estudo, denominam histórico de doação como o número de ciclos finalizados anteriormente.

³ Whyte, nesse estudo, também denominou histórico de doação como o número de ciclos finalizados anteriormente.

⁴ Flegel, Besenfelder e Wagner, neste estudo, sugeriram um *score* baseado no número de doações anteriores e quão recente foi a última doação, o qual chamaram de histórico de doação.

doadores de sangue residentes em Land, estado de Baden-Wurttemberg, Alemanha.

Apesar desses indicativos relacionados com o ato de doar, sabemos que muitos fatores envolvem ainda a decisão de doar sangue. A disponibilidade, o acesso aos postos de coleta entre outros motivos, são aspectos que podem influenciar nesta decisão.

Além disso, algumas situações especiais podem ampliar a incidência de doações de primodoadores, mas não necessariamente os convertem em doadores regulares. Um exemplo disso é o aumento de primodoadores logo após os ataques terroristas de 11 de setembro de 2001, nos Estados Unidos, sem um conseqüente aumento no risco absoluto de infecções virais decorrentes de transfusões. Porém, após os ataques terroristas, as taxas de retorno destes doadores foram tão baixas quanto aquelas que já vinham sendo observadas antes da ocorrência dos mesmos (GLYNN, et al 2003).

Por isso, não é difícil inferir a presença de um subgrupo de doadores que após iniciarem o seu ciclo de doação, não irão retornar para uma subsequente doação. Tal proporção de doadores improváveis de retornar, no contexto de Análise de Sobrevida, pode ser estimada via Modelos de Longa Duração, uma classe de modelos que incorpora o fato de existirem indivíduos “imunes” ao evento de interesse, no caso, o retorno.

Com o intuito de oferecer um conhecimento aprimorado do comportamento de retorno do doador que inicia seu ciclo de doação voluntariamente, o presente estudo objetiva quantificar esse comportamento utilizando técnicas de Análise de Sobrevida utilizando os registros do Centro Regional de Hemoterapia do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo (HCFMRP – USP) e associá-lo com algumas variáveis sócio-demográficas (covariáveis) já apontadas na literatura como sendo um diferencial para a segmentação de subgrupos mais propensos ou menos propensos à retornar.

Saber se um determinado doador irá ou não se submeter a uma próxima doação é um importante desafio das agências coletoras de sangue, e seu impacto está diretamente

relacionado à garantia do fornecimento sustentável da demanda por bolsas de sangue. Identificar um modelo estatístico capaz de descrever este processo poderá ser um instrumento útil para entender melhor esse público e como lidar com ele.

1.1 As Políticas de Sangue no Brasil

A política de sangue e hemoderivados no Brasil até 1980 era crítica, caracterizada por um sistema desorganizado e desigual na qualidade dos serviços prestados, carecia de uma rígida fiscalização. Com o advento da AIDS, o tema ganhou destaque principalmente pelo número de casos transmitidos por transfusão sanguínea, cerca de 2% dos casos (JUNQUEIRA; ROSENBLIT; HAMERSCHLAK, 2005), chegando a 8,8% o número de casos notificados pelo Ministério da Saúde de sua transmissão e de outras doenças transmitidas através do sangue (BRASIL, 2002a).

Na região de Ribeirão Preto, a situação dos serviços hemoterápicos não era diferente. Tal situação decorreu devido à maioria dos bancos de sangue apresentar irregularidades tais como a deficiência na realização de exames sorológicos, triagem inadequada dos doadores, condições impróprias para o armazenamento do sangue, área física inadequada, entre outros⁵.

A partir daí, na década de 80, foram tomadas medidas rigorosas no sentido de oferecer maior segurança aos doadores e receptores de sangue e hemoderivados, entre elas, a criação da Política Nacional do Sangue, a campanha da doação altruísta de sangue e a formulação da constituição de 1980 (JUNQUEIRA; ROSENBLIT; HAMERSCHLAK, 2005).

⁵ Histórico – Centro Regional de Hemoterapia de Ribeirão Preto. Política Nacional para o setor de hemoterapia. Pasta (1). Agosto de 1987 a Dezembro de 1990

No Estado de São Paulo, através das Secretarias da Saúde, foi criado em 1987 o GEPRO (Grupo Especial do Programa de Hematologia/Hemoterapia) com o intuito de desenvolver uma política estadual para o setor. Uma das suas propostas foi à implantação de um sistema descentralizado de serviços, com a instituição de um Hemocentro de “referência” em cada uma das quatro regiões de saúde do interior do estado (Campinas, Botucatu, Marília e Ribeirão Preto), que deveriam assumir responsabilidade pela execução das políticas relativas ao sangue em sua respectiva área de abrangência⁶.

Ainda visando à melhoria na qualidade da oferta de sangue e hemoderivados, em 1998, o Ministério da Saúde instituiu o Programa Nacional de Doação Voluntária de Sangue (PNDVS) como parte da Meta Mobilizadora Nacional (“Sangue - 100% com garantia de qualidade em todo o seu processo até 2003”), onde o seu principal objetivo era envolver a sociedade no processo de doação de sangue através da sua participação consciente e responsável, e com isso garantir a quantidade adequada à demanda do país e a melhoria da sua qualidade (BRASIL, 1998).

Para orientar as ações a serem tomadas na fidelização da doação, foi realizada em 2004, uma pesquisa em vários hemocentros do Brasil patrocinada pela Agência Nacional de Vigilância Sanitária (ANVISA), por solicitação do PNDVS, onde traçou um quadro demonstrativo das características pessoais, opiniões, atitudes e hábitos expressos pelos doadores de sangue, permitindo conhecer a nível regional e nacional um pouco mais desse público (BRASIL, 2004c).

Assim, do conjunto de pessoas consultadas, houve uma maior prevalência do sexo masculino (62,4%), de faixa etária de 30 a 39 anos (28,2%), estado civil solteiro (49,0%), etnia branca (49,4%), nível de escolaridade ensino médio completo (39,7%) e classe

⁶ Histórico – Centro Regional de Hemoterapia de Ribeirão Preto. Política Nacional para o setor de hemoterapia. Pasta (1). Agosto de 1987 a Dezembro de 1990.

sócio-econômica C (42,9%). Também nota-se que entre a maioria dos respondentes não existe outro doador na família (52,3%).

Com relação aos fatores comportamentais, verificou-se que 69,6% dos entrevistados já tinham doado sangue anteriormente e destes, a maioria havia doado nos últimos 13 meses (67,3%). Também observou-se que 53,5% dos indivíduos disseram ter doado cinco vezes ou mais em sua vida. A periodicidade de doação de maior incidência foi de três em três meses (19,4%).

Com relação aos fatores motivacionais, 70,2% informaram que a doação foi espontânea / voluntária e a principal razão ter doado era para ajudar o próximo / parente / amigo (36,4%). O principal motivo por não ter doado no último ano foi a falta de tempo, (23,8%), sendo interessante chamar a atenção para outros fatores citados como: ninguém pediu (7,8%), doença (7,4%), falta de interesse (6,0%), falta de oportunidade (5,4%) e distância (4,3%).

Na opinião geral dos entrevistados, os motivos pelos quais as pessoas não doam sangue são o medo (35,0%) e a falta de informação (18,6%) e que deveria haver mais campanhas para incentivar as pessoas a doarem sangue (23,1%) e uma maior divulgação (14,5%). Ver ou ouvir mensagens também é útil não só para conscientizar as pessoas, mas também para lembrá-las (28,1%).

Com relação ao Serviço de Saúde, a maioria dos entrevistados disse ter doado pela última vez numa Instituição Pública, (87,1%), e que em geral a qualidade do serviço de hemoterapia é considerado entre boa a excelente, (96,3%). O principal motivo desta avaliação positiva é o bom atendimento (50,3%), seguido pela higiene (9,0%) e pela rapidez (7,8%). Entre os doadores, 5,7% estavam insatisfeitos com a doação de sangue, sendo que o principal motivo da insatisfação foi o mau atendimento (cerca de 23,5% destes).

Estima-se que no Brasil, em torno de três milhões de bolsas de sangue são coletadas, o que corresponde a 1,7% da população brasileira (BRASIL, 2002b). O intuito é elevar o percentual de doadores para 2,2% da população em 2005 e atingir 3,0% até 2007, o que corresponde ao nível recomendado pela Organização Mundial de Saúde (OMS) e o limite que alcança mais segurança (BRASIL, 2004b).

1.2 O Centro Regional de Hemoterapia de Ribeirão Preto

O Centro Regional de Hemoterapia de Ribeirão Preto, responsável pela coleta, processamento, distribuição e transfusão de frações e derivados sanguíneos, é constituído além da sua unidade central, o Hemocentro de Ribeirão Preto, por uma rede de quatro Núcleos de Hematologia e Hemoterapia - Araçatuba, Fernandópolis, Franca e Presidente Prudente, quatro Unidades de Hemoterapia - Batatais, Bebedouro, Olímpia e Serrana, um Posto de Coleta em Ribeirão Preto, além de várias agências transfusionais.

Sua região de abrangência é composta de 187 municípios com aproximadamente 4 milhões de habitantes. Possui um banco de dados com cerca de 158,9 mil doadores aptos cadastrados em toda a região, obtendo uma média de 8000 doações/mês.

Fundado como parte da ação do governo face a situação em que se encontrava o país frente ao setor, constitui hoje um centro de referência na coleta, processamento e distribuição de sangue e hemocomponentes, e certificado pela ISO (NBR 9001:2000) e acreditado pela AABB – *American Association of Blood Bank*.

Informações disponíveis no endereço eletrônico (<http://pegasus.fmrp.usp.br>) apresentam um breve histórico do centro.

A captação de doadores atua em duas frentes de coleta: coleta interna e coleta externa com o objetivo de desenvolver campanhas de doação de sangue junto à população.

1.3 A doação de Sangue: conceitos e definições

No Brasil, a doação de sangue deve ser um ato voluntário, assim como disposto na Constituição da República por meio da Resolução nº.343/2002 (BRASIL, 2002c) onde estabelece que não seja admitido qualquer tipo de remuneração para a sua realização (GUERRA, 2005).

Existem vários tipos de doação de sangue, de acordo com a motivação do doador e de tipos de doadores, e de acordo com a periodicidade da doação. A ANVISA, por meio da Resolução nº149/2001 (BRASIL, 2001), define os tipos de doação de sangue como:

- ***Doação Autóloga:*** é aquela doação advinda do indivíduo que doa para si próprio. Este procedimento só pode ser feito por indicação e critério médico.
- ***Doação Espontânea:*** é aquela doação advinda do indivíduo que procura o serviço de hemoterapia espontaneamente, independente de campanha, atendimento a um paciente específico ou para atender a uma convocação da necessidade do serviço;
- ***Doação de Reposição:*** é aquela doação advinda do indivíduo que doa para atender a necessidade de um paciente, inclusive as doações dirigidas, vinculadas ou específicas. A doação vinculada é destinada à reposição do estoque de sangue no Hemocentro e realizada por um amigo ou parente do paciente que recebeu ou irá receber transfusão. Este tipo de doação não é obrigatório. A doação específica é dirigida para um determinado paciente que requer cuidado especial por causa de eventuais reações adversas no momento da transfusão.

Com relação ao tipo de doador, a ANVISA classifica como:

- ***Doador de Primeira vez:*** é aquele indivíduo que doa pela primeira vez naquele serviço de Hemoterapia.
- ***Doador Esporádico:*** é aquele indivíduo que doa eventualmente num intervalo superior a treze meses.
- ***Doador de Repetição:*** é aquele indivíduo que doa pelo menos uma vez a cada treze meses, considerando a data de sua última doação no mesmo Serviço de Hemoterapia. Também considerado doador de retorno, fidelizado ou habitual, dependendo do número de doações/ano.

Quanto ao procedimento da doação realizado, há dois tipos:

- ***Doação Fracionária:*** a doação do sangue é total, sendo fracionada posteriormente nos diversos hemocomponentes do sangue para o seu melhor aproveitamento.
- ***Doação por Aféreses:*** significa um procedimento com separação de algum hemocomponente do sangue através de um equipamento: plaquetas, hemáceas, plasma, granulócitos e linfócitos. Trata-se de um procedimento especial que envolve técnicas apuradas. É realizado com o auxílio de um equipamento específico e monitorado por médicos e enfermeiros.

CAPÍTULO II

Objetivos

2.1 Geral

Introduzir modelos de análise de sobrevivência voltados à descrição do comportamento de retorno dos doadores de sangue que compareceram pelo menos uma vez em algum serviço de hemoterapia localizado na região de abrangência do Centro Regional de Hemoterapia de Ribeirão Preto.

2.2 Específicos

- 1- Descrever o histórico de doação e de doadores da região estudada.
- 2- Propor modelos estatísticos que permitam estudar o tempo entre a primeira e a segunda doação do doador de sangue voluntário, bem como associá-lo com algumas variáveis sócio-demográficas, buscando identificar subgrupos de indivíduos propensos a retornar, bem como estimar a proporção de indivíduos que jamais retornarão.

O presente estudo envolve a construção de modelos probabilísticos aplicados no contexto de doação de sangue utilizando as informações registradas no Centro Regional de Hemoterapia do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, HCFMRP – USP. O estudo foi conduzido a partir de um levantamento retrospectivo do histórico de doação de todos os indivíduos que compareceram para uma doação no período de 23/07/1996⁷ até 23/06/2005 em algum serviço de hemoterapia da região de abrangência estudada (Ver Anexo A). Também foram obtidas informações referentes às características sócio-demográficas, comportamentais e motivacionais do doador. Este estudo foi autorizado pela direção do Hemocentro (Anexo B) e aprovado pelo Comitê de Ética em Pesquisa do HCFMRP – USP (Anexo C).

3.1 Casuística

Compõem o banco de dados registros de todos os primodoadores de sangue voluntários⁸ que compareceram em alguma Unidade, Núcleo ou Hemocentro (coleta interna) pertencente à região de abrangência do Centro Regional de Hemoterapia de Ribeirão Preto ou compareceram em uma unidade móvel de coleta de sangue (coleta externa) na mesma região no período de 23/07/1996 a 23/06/2005 e que não se incluíram nos critérios de exclusão.

3.2 Critérios de Inclusão/Exclusão

Foram excluídos os registros de doações ou doadores que satisfizeram os itens:

⁷ A partir desta data, iniciou o funcionamento do sistema informatizado (Sistema de Controle de Doadores e Laboratórios), utilizado atualmente no Hemocentro. Nesse caso, consideramos o primeiro registro encontrado na base de dados o início do ciclo de doação de um determinado doador, sendo, portanto, a primeira vez que aquele indivíduo doou naquele serviço de hemoterapia a partir de 23/07/1996.

⁸ Doador que iniciou o seu ciclo de doação pela primeira vez naquele serviço de hemoterapia a partir daquela data em caráter voluntário. A partir daqui, denominaremos simplesmente por doador de sangue.

- Ter sido reprovado definitivamente na primeira doação de sangue: a condição de reprovado definitivamente era vinculado à etapa de triagem. Assim, dado que o levantamento realizado corresponde a um período extremamente longo, critérios de triagem antes considerados determinantes para classificar um determinado doador impossibilitado de doar não necessariamente correspondem aos mesmos critérios adotados hoje. Por isso, casos em que um doador classificado como reprovado definitivamente retornou a doar podem ser encontrados. Dessa maneira, foram excluídos apenas os primodoadores inaptos definitivamente por considerar um fenômeno propício ao não retorno.
- Não ter a condição de aprovado registrada no banco de dados nas sucessivas doações: é muito provável que o fato de não ter a condição de aprovado no banco de dados venha a ser caracterizado como uma não doação, ou seja, o indivíduo foi até o local de coleta, se cadastrou, mas não completou o ciclo de doação.
- Ter realizado um Exame Extra Rotina⁹: esse tipo de procedimento não caracteriza efetivamente uma intenção de doar sangue e por isso esse público foi desconsiderado.
- Ter realizado uma doação em Altinópolis: a captação de doadores nesta cidade foi pouca expressiva, além de ser um local já extinto do programa de doação como um posto de coleta.
- Ter realizado uma doação autóloga: esse tipo de doação caracteriza um público com características particulares que diferem muito do comportamento dos outros tipos de doação.

⁹ Doador que passa por todo o processo de triagem que antecede a doação mais não doa efetivamente uma bolsa de sangue; apenas colhe amostras para exames e aguarda um momento específico para doar, geralmente por convocação do serviço de hemoterapia.

- Ter realizado uma doação por Aféreses: esse tipo de procedimento de doação também caracteriza um público com características particulares.

Foram incluídos os registros de doações no qual os seguintes itens foram satisfeitos:

- Ter entre 18 a 65 anos de idade: essa faixa etária é considerada permissível à doação. Menores de 18 anos e maiores de 65 anos caracterizam casos especiais que requer muitas vezes autorização do responsável no primeiro caso e por isso foram desconsiderados.
- Ter dado tempo suficiente de retornar para doar sangue de acordo com o período que data o fim do seguimento no estudo: assim, indivíduos do sexo masculino que doaram pela primeira vez após o dia 23/04/2005 e indivíduos do sexo feminino que doaram pela primeira vez após o dia 23/03/2005 foram desconsiderados.

3.3 Variáveis

Variáveis Dependentes:

- (a) Tempo decorrido entre o primeiro e o segundo ciclo de doação (para os indivíduos que retornaram) ou tempo até o fim do seguimento (para os indivíduos que não retornaram) - em dias¹⁰.
- (b) Número de retornos¹¹ do doador – esta variável assume zero caso um doador não apresente registros de doações.

¹⁰ A partir daqui, denominaremos esse tempo como tempo até o retorno do doador, ou tempo entre a primeira e a segunda doação.

¹¹ Definiremos retorno do doador como o número de ciclos de doação finalizados correspondente ao número de registros encontrados na base de dados para um determinado doador durante o período de seguimento menos a primeira doação.

Variáveis Independentes:

(a) Variáveis Comportamentais:

- Tipo de Doador: foram consideradas as categorias: doador de primeira vez, doador esporádico e doador de repetição.
- Condição da Doação: foram consideradas as categorias: aprovado, inapto temporariamente, inapto definitivamente e não informado.

(b) Variáveis Motivacionais:

- Tipo de Doação: correspondente na primeira doação registrada, sendo considerada as categorias: autóloga, espontânea (voluntária), reposição (específica ou vinculada) e não informado.

(c) Variáveis Sócio-Demográficas fixas:

- Cor da pele: registrada e avaliada pela atendente, sendo considerada as categorias: amarela, branca, morena ou parda, negra ou mulata e não informado.
- Estado Civil: foram consideradas as categorias: casado, divorciado, outros, solteiro, viúvo e não informado.
- Nível de escolaridade: sendo considerada as categorias: analfabeto, primeiro grau incompleto, primeiro grau completo, segundo grau incompleto, segundo grau completo, terceiro grau incompleto, terceiro grau completo, mestrado, doutorado, pós graduação e não informado.
- Tipo Sanguíneo (de acordo com o sistema ABO) e respectivo fator RH: sendo considerada as categorias: A-, A+, B-, B+, AB+, AB-, O-, O+ e não informado ou não obtido.
- Sexo: foram consideradas as categorias: feminino e masculino.

(c) Variáveis Sócio-Demográficas que variam com os ciclos de doação:

- Faixa Etária: sendo considerada as categorias: 18 a 19 anos, 20 a 29 anos, 30 a 39 anos, 40 a 49 anos, 50 a 59 anos, 60 anos ou mais e não informado.
- Unidade de Coleta: categorizada a partir da variável local de coleta de acordo com o número de doadores cadastrados, sendo consideradas as categorias: Unidades Pequenas (Unidades de Batatais, Bebedouro, Olímpia e Serrana), Unidades Grandes (Núcleos de Araçatuba, Fernandópolis, Franca e Presidente Prudente, Hemocentro de Ribeirão Preto e Posto de Coleta Quintino em Ribeirão Preto) e Coleta Externa (Ribeirão Preto, Araçatuba, Fernandópolis e Franca).

3.4 Metodologia

No presente estudo, serão utilizadas as seguintes ferramentas estatísticas:

1. Para descrever o comportamento do tempo entre a primeira e a segunda doação será utilizado o estimador de Kaplan-Meier. Através dele, podemos estimar a função de sobrevivência e consequentemente obter uma medida da propensão de não retornar dos doadores em todo o período de seguimento. Essa medida também pode ser interpretada como a “chance” do doador retornar para uma subsequente doação após um determinado período de acompanhamento. Também será estimado o risco instantâneo do doador retornar ao longo do período de doação através da estimação da função risco. Tais estimativas serão obtidas para cada uma das variáveis sócio-demográficas (covariáveis).

2. Para estudar o comportamento do tempo entre a primeira e a segunda doação e o impacto das variáveis sócio-demográficas serão utilizados:

- Modelos de Sobrevida Paramétricos.
- Modelos de Sobrevida Paramétricos de Longa-Duração.

3.4.1 A Análise de Sobrevida

Como foi visto, algumas ferramentas de Análise de Sobrevida são amplamente utilizadas no estudo do comportamento de retorno do doador de sangue: James e Matthews (1993, 1996) utilizaram os estimadores de Kaplan-Meier para destacar que o intervalo entre as doações está inversamente relacionado ao número de doações anteriores e para evidenciar fatores que discriminam subgrupos de doadores mais propensos e menos propensos a retornar. Além disso, citam a função risco como um indicador para verificar como o comportamento do doador varia ao longo do tempo. Ownby, Watanabe e Nass (1999), também utilizaram esses estimadores para destacar alguns fatores sócio-demográficos possivelmente associados com o retorno à doação. Melnyk et al. (1995), fazem uma revisão geral do uso destas ferramentas no contexto de doação de sangue.

A Análise de Sobrevida (HOUGAARD, 1999; COLOSIMO; GIOLO, 2006; LOUZADA-NETO; MAZUCHELI; ACHCAR, 2002) consiste de um conjunto de procedimentos para a análise estatística de dados cuja variável resposta é o tempo até a ocorrência de um evento de interesse, determinado a partir de um tempo inicial pré-estabelecido. Aqui, o tempo estudado é aquele medido entre o primeiro registro de doação até o seu retorno para uma subsequente doação, sendo denominado genericamente de tempo de sobrevida.

Sua principal característica é a capacidade de extrair informações de dados censurados, ou seja, daqueles indivíduos dos quais não foi observada a ocorrência do evento até o fim do período estudado. No caso, corresponde aos doadores que não foram observados o retorno para uma subsequente doação até o fim do seguimento.

Na Análise de Sobrevida, o comportamento da variável aleatória tempo de sobrevida, $T \geq 0$, pode ser expresso através de várias funções, como: a função densidade de probabilidade, $f(t)$, a função de sobrevivência, $S(t)$ e a função risco, $h(t)$. Cada função descreve um determinado aspecto da amostra.

A função densidade de probabilidade é definida como o limite da probabilidade de observar o evento de interesse em um indivíduo no intervalo de tempo $[t, t + \Delta t]$ por unidade de tempo, expressa por,

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}, \quad (3.1)$$

onde $f(t) \geq 0$, para todo t , e tem a área abaixo da curva igual a 1.

A função de sobrevivência é definida como a probabilidade de um indivíduo “sobreviver” até o tempo t , ou seja, a probabilidade de ocorrer o evento além de t , sendo definida por,

$$S(t) = P(T > t) = 1 - F(t), \quad (3.2)$$

tal que $S(t) = 1$ quando $t = 0$ e $S(t) = 0$ quando $t \rightarrow \infty$ e $F(t) = \int_0^t f(u) du$ representa a função distribuição acumulada.

A função Risco, ou taxa de falha instantânea, é definida como o limite da probabilidade de observar o evento de interesse para um indivíduo no intervalo de tempo $[t, t + \Delta t]$ dado que o mesmo tenha “sobrevivido” até o tempo t , dada por,

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}. \quad (3.3)$$

Esta função também pode ser definida utilizando (3.1) e (3.2) por meio da expressão,

$$h(t) = \frac{f(t)}{S(t)}, \quad (3.4)$$

onde $h(t) \geq 0$, e representa o relacionamento entre as três funções utilizadas para descrever os dados em Análise de Sobrevidência.

A função densidade de probabilidade também é definida como a derivada da função densidade acumulada, ou seja,

$$f(t) = \frac{\partial F(t)}{\partial t}. \quad (3.5)$$

Como $F(t) = 1 - S(t)$, podemos escrever,

$$f(t) = \frac{\partial [1 - S(t)]}{\partial t} = -S'(t). \quad (3.6)$$

Substituindo (3.6) em (3.4), temos,

$$h(t) = -\frac{S'(t)}{S(t)} = -\frac{\partial [\log S(t)]}{\partial t}, \quad (3.7)$$

Dessa forma, obtemos,

$$\log S(t) = -\int_0^t h(u) \partial u, \quad (3.8)$$

ou seja,

$$S(t) = \exp\left(-\int_0^t h(u) \partial u\right). \quad (3.9)$$

Uma outra função que pode ser obtida de (3.3) é dada pela função risco acumulada, dada por,

$$H(t) = \int_0^t h(u) du. \quad (3.10)$$

Substituindo (3.10) em (3.9), temos que,

$$S(t) = \exp[-H(t)]. \quad (3.11)$$

3.4.1.1 Estimador Não Paramétrico de Kaplan-Meier

Como o interesse está voltado em estimar as funções descritas acima, um estimador não paramétrico muito conhecido na literatura é o de Kaplan-Meier (KAPLAN; MEIER, 1958) o qual permite estimar as funções descritas acima considerando a presença de observações censuradas.

Considere neste estudo n indivíduos, e seus respectivos tempos de sobrevida, incluindo as censuras, ordenados, isto é, $t_1 \leq t_2 \leq \dots \leq t_n$. A função de sobrevivência empírica obtida a partir do estimador de Kaplan-Meier é dada por,

$$\begin{aligned} \hat{S}_{KM}(t) &= \frac{n_1 - \delta_1}{n_1} \frac{n_2 - \delta_2}{n_2} \dots \frac{n_r - \delta_r}{n_r}, \\ &= \prod_{r: t_r \leq t} \frac{n_i - \delta_i}{n_i}, \end{aligned} \quad (3.12)$$

onde t_r é o maior tempo de sobrevida menor ou igual a t , n_i é o número de indivíduos sob risco (indivíduos que não ocorreram o evento e não foram censurados) em t_i e δ_i representa o número de ocorrências do evento no tempo t_i ($\delta_i = 0$ quando for observado a não ocorrência do evento - tempo de sobrevida censurado; e $\delta_i = 1$ quando for observado a ocorrência do evento - tempo de sobrevida não censurado), onde i pode ser qualquer valor inteiro entre 1 e r .

O estimador de Kaplan-Meier da função risco no intervalo de tempo $[t, t + \Delta t]$

é dado por,

$$\hat{h}_{KM}(t) = \frac{1}{\Delta t} \left(1 - \frac{n_i - \delta_i}{n_i} \right). \quad (3.13)$$

Enquanto que a estimação de (3.12) nos informa a probabilidade de um doador retornar após t dias de seguimento, a estimação de (3.13) nos mostra como a chance pontual (não acumulada) de um doador retornar se modifica com o passar do tempo de acompanhamento, identificando os períodos de maior “risco” deste retornar.

Tais funções podem ser segmentadas de acordo com uma determinada covariável de interesse e assim exibir o seu comportamento de acordo com cada categoria. Porém, quando o interesse está em identificar a relação e a influência de uma ou mais covariáveis nos tempos de sobrevida, a utilização de modelos de regressão pode ser mais apropriada como será visto a seguir.

3.4.1.2 Modelos Paramétricos em Análise de Sobrevivência

A construção de modelos em Análise de Sobrevivência busca explorar e conhecer como o tempo de sobrevida está relacionado com uma ou mais covariáveis de interesse. Nesse caso, podemos identificar quais covariáveis estão mais associadas com o evento estudado bem como a intensidade e a direção de cada uma delas em explicar a ocorrência do evento (HOUGAARD, 1999; COLOSIMO; GIOLO, 2006, p. ; LOUZADANETO; MAZUCHELI; ACHCAR, 2002). No caso, o intuito é estudar o tempo que o doador leva para retornar para sua segunda doação, de acordo com as variáveis relacionadas com o seu perfil de tal maneira que possamos encontrar evidências de subgrupos de doadores mais ou menos propensos a retornar. Para isso, existem duas classes de modelos propostas na literatura: os Modelos Paramétricos e os Semi-Paramétricos.

Enquanto os Modelos Paramétricos assumem uma distribuição conhecida de probabilidade para os tempos de sobrevida os Modelos Semi-Paramétricos, também chamados de Modelos de Regressão de Cox (COX, 1972), têm, como característica principal, o pressuposto de proporcionalidade dos riscos entre as categorias de uma determinada covariável (ver, por exemplo, LEE; WANG, 2003).

No contexto de doação de sangue, a modelagem do tempo até uma subsequente doação incorporando covariáveis, geralmente é baseado nos modelos de Cox, como visto em Melnyk et al. (1995), James e Mattheus (1996) e Ownby, Watanabe e Nass (1999). Porém, não é sempre que a suposição de proporcionalidade dos riscos entre as categorias de uma determinada covariável é atendida.

Uma alternativa é construir modelos em diferentes períodos de tempo, de tal maneira que em cada período essa suposição seja considerada satisfeita, como sugerida em James e Mattheus (1996). Outras maneiras de lidar com a não proporcionalidade dos riscos podem ser vista em Schemper (1992).

Porém, quando conhecemos a distribuição dos tempos de sobrevida, o ajuste de um modelo paramétrico pode trazer mais informações sobre a natureza da sua distribuição e do comportamento da função risco ao longo do tempo. Além disso, é um modelo mais flexível, dado a sua facilidade em incorporar o efeito das covariáveis em seus parâmetros, como será visto neste capítulo.

Algumas distribuições mais utilizadas para descrever os tempos de sobrevida são: Exponencial, Weibull, Log-Normal, Log-Logística, entre outras.

A distribuição Exponencial acomoda funções de risco constante, enquanto que a distribuição de Weibull acomoda funções monótonas, isto é, crescente, decrescente ou constante. As distribuições Log-Logística e Log-Normal acomodam funções de risco unimodais (LOUZADA-NETO; MAZUCHELI; ACHCAR, 2002).

Sendo intuitivo afirmar que quanto maior o tempo entre as doações menor é a chance do indivíduo retornar, a distribuição de Weibull pode ser apropriada para estudar o comportamento de retorno do doador, assumindo assim uma função risco decrescente. Porém, como não descartamos a possibilidade de uma grande parte dos doadores retornarem em algum momento, mas não necessariamente logo após o intervalo mínimo entre as doações¹², é provável a existência de um “pico” na função risco em um determinado momento. Nesse caso as distribuições Log-Normal e Log-Logística podem ser mais apropriadas.

O comportamento do tempo entre a primeira e a segunda doação assumindo seguir uma distribuição Weibull, Log-Normal, Log-Logística serão descritas a seguir (ver, por exemplo, LEE; WANG, 2003). Também discutiremos o caso da Gama Generalizada por ser uma distribuição mais flexível que as anteriores (YAMAGUCHI, 1992; PENG; DEAR; DENHAM, 1998).

A distribuição Weibull

Considere t uma observação de uma variável aleatória T , representando o tempo até o retorno do doador de sangue ou até o seu não retorno até o fim do seguimento seguindo uma distribuição Weibull, cuja função densidade de probabilidade é dada por,

$$f(t) = \lambda \gamma (\lambda t)^{\gamma-1} e^{-(\lambda t)^\gamma}, \quad (3.14)$$

onde $t \geq 0$, $\lambda > 0$ é o parâmetro de locação e $\gamma > 0$ é o parâmetro de escala.

Dada a relação existente em (3.4), a função de sobrevivência e sua respectiva função risco, são dadas por,

$$S(t) = e^{-(\lambda t)^\gamma}, \quad (3.15)$$

¹² Tempo mínimo necessário para que o organismo se recupere da doação anterior, sendo estipulado por aproximadamente 60 dias para homens e 90 dias para mulheres.

e

$$h(t) = \lambda \gamma (\lambda t)^{\gamma-1}, \quad (3.16)$$

sendo uma função monótona estritamente crescente para $\gamma > 1$, estritamente decrescente para $\gamma < 1$ e constante para $\gamma = 1$. Para $\gamma = 1$ tem-se a função risco da distribuição Exponencial, que é um caso particular da distribuição Weibull.

A distribuição Log-Normal

Se T é uma variável aleatória, tal que $\ln(T)$ tem distribuição normal com média μ e variância σ^2 , então dizemos que T tem distribuição Log-Normal, com função de densidade de probabilidade dada por,

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\left[-\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma}\right)^2\right], \quad (3.17)$$

onde $t > 0$, μ e σ são respectivamente a média e o desvio padrão de T quando esta variável é tomada em logaritmos.

As funções de sobrevivência e função risco neste caso, não apresentam uma forma analítica explícita, sendo expressa por,

$$\begin{aligned} S(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \frac{1}{x} \exp\left[-\frac{1}{2\sigma^2}(\ln x - \mu)^2\right] dx \\ &= 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \end{aligned}, \quad (3.18)$$

onde $\Phi(\cdot)$ é a função distribuição acumulada da distribuição normal.

A partir de (3.17) e (3.18), a função risco pode ser expressa utilizando a expressão (3.4) por,

$$h(t) = \frac{\left(1/\left(\sqrt{2\pi}\sigma t\right)\right) \exp\left[-\frac{1}{2}\left((\ln t - \mu)/\sigma\right)^2\right]}{1 - \Phi\left((\ln t - \mu)/\sigma\right)}, \quad (3.19)$$

sendo uma função não monótona, ou seja, ela cresce, atinge um valor máximo e depois decresce.

Seja $\lambda = \exp(-\mu)$. Então $-\mu = \ln \lambda$. Assim, podemos reescrever (3.17) e (3.18) respectivamente como,

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\left[-\frac{1}{2\sigma^2}(\ln(\lambda t))^2\right], \quad (3.20)$$

e

$$\begin{aligned} S(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_x^\infty \frac{1}{x} \exp\left[-\frac{1}{2\sigma^2}(\ln(\lambda x))^2\right] dx \\ &= 1 - \Phi\left(\frac{\ln(\lambda t)}{\sigma}\right), \end{aligned} \quad (3.21)$$

onde $\Phi(\cdot)$ é a função distribuição acumulada da distribuição normal.

A distribuição Log-Logística

Se T é uma variável aleatória, tal que $\ln(T)$ tem distribuição logística, então dizemos que T tem distribuição Log-Logística, com função de densidade de probabilidade dada por,

$$f(t) = \frac{\lambda \gamma t^{\gamma-1}}{(1 + \lambda t^\gamma)^2}, \quad (3.22)$$

onde $t \geq 0$, $\lambda > 0$ é o parâmetro de locação e $\gamma > 0$ é o parâmetro de escala.

Dada a relação existente em (3.4), a função de sobrevivência e sua respectiva função risco, são dadas por,

$$S(t) = \frac{1}{(1 + \lambda t^\gamma)}, \quad (3.23)$$

e

$$h(t) = \frac{\lambda \gamma t^{\gamma-1}}{(1 + \lambda t^\gamma)}, \quad (3.24)$$

onde, para $\gamma > 1$, o risco é crescente alcançando um pico e a partir daí começa a declinar, similar a função risco da distribuição log-normal; para $\gamma < 1$, o risco é decrescente, similar a função risco da distribuição Weibull; para $\gamma = 1$, o risco também é decrescente.

A distribuição Gama e Gama Generalizada

Seja T uma variável aleatória com distribuição Gama, com função densidade de probabilidade dada por,

$$f(t) = \frac{\lambda}{\Gamma(\alpha)} (\lambda t)^{\alpha-1} e^{-\lambda t}, \quad (3.25)$$

onde $t > 0, \lambda > 0$ é o parâmetro de locação, $\alpha > 0$ é o parâmetro de forma e $\Gamma(\cdot)$ é a função

gama completa definida por $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx = (\alpha-1)!$.

A função de sobrevivência apresenta uma forma complexa, dada por,

$$S(t) = \int_t^\infty \frac{\lambda}{\Gamma(\alpha)} (\lambda x)^{\alpha-1} e^{-\lambda x} dx. \quad (3.26)$$

A função risco, pode ser obtida a partir de (3.4) e apresenta um padrão crescente de 0 até α quando t cresce de 0 a infinito, se $\alpha > 1$, o padrão é decrescente de infinito até α quando t cresce de 0 a infinito, se $0 < \alpha < 1$, e é constante, se $\alpha = 1$. Para $\alpha = 1$ tem-se a função risco da distribuição Exponencial, que é um caso particular da distribuição Gama.

Outra distribuição que merece destaque é a distribuição Gama Generalizada, caracterizada por três parâmetros, onde sua função densidade de probabilidade é definida como,

$$f(t) = \frac{\gamma \lambda^{\alpha\gamma}}{\Gamma(\alpha)} t^{\alpha\gamma-1} e^{-(\lambda t)^\gamma}, \quad (3.27)$$

onde $t > 0, \lambda > 0$ é o parâmetro de locação, $\gamma > 0$ é o parâmetro de escala e $\alpha > 0$ é o parâmetro de forma o que a torna bastante flexível. Note que a partir dela, encontramos alguns casos especiais, tal que, se $\alpha = \gamma = 1$, então $T \sim \text{Exp}(\lambda)$; se $\alpha = 1$, então $T \sim \text{Weibull}(\lambda, \gamma)$; se $\gamma = 1$, então $T \sim \text{Gama}(\lambda, \alpha)$. Além disso, a distribuição Log-Normal aparece como um caso limite da distribuição gama generalizada quando $\alpha \rightarrow \infty$.

A função (3.27) também pode ser definida como tendo a seguinte função densidade de probabilidade,

$$f(t) = \frac{\gamma / \alpha^\alpha \lambda^{\alpha\gamma}}{\Gamma(\alpha)} t^{\alpha\gamma-1} \exp[-\alpha(\lambda t)^\gamma], \quad (3.28)$$

onde $t > 0, \lambda > 0$ e $\alpha > 0$, comumente referida como distribuição Gama Generalizada Estendida (GGE) e pode ser mostrado que se $\alpha = \gamma = 1$, então $T \sim \text{Exp}(\lambda)$; se $\alpha = 1$ então $T \sim \text{Weibull}(\lambda, \gamma)$; se $\gamma = 1$, então $T \sim \text{Gama}(\lambda, \alpha)$, além da distribuição Log-Normal quando $\alpha \rightarrow \infty$ (Figura 3.4.1).

As funções de sobrevivência e função risco neste caso, não apresentam uma forma analítica explícita, sendo expressa por,

$$S(t) = \begin{cases} I(\alpha(\lambda t)^\gamma, \alpha) & \text{se } \gamma < 0 \\ 1 - I(\alpha(\lambda t)^\gamma, \alpha) & \text{se } \gamma > 0 \end{cases}, \quad (3.29)$$

onde $I(a, x)$ é a função gama incompleta definida por $I(a, \alpha) = \frac{1}{\Gamma(\alpha)} \int_0^a x^{\alpha-1} e^{-x} dx$ e a função risco pode ser obtida através de (3.4).

Figura 3.4.1. Casos particulares da distribuição Gama Generalizada Estendida (GGE).

O comportamento da função de sobrevivência assumindo uma distribuição GGE para o tempo até o retorno do doador, dada em (3.29), atribuindo diferentes valores para seus parâmetros pode ser visto na Figura 3.4.2.. Note que valores de λ grandes caracterizam um decrescimento mais rápido de $S(t)$. A forma da função de sobrevivência é determinada pelo parâmetro α , onde, quanto maior o seu valor, mais acentuado tende a ser a curva em formato de “S”. O parâmetro γ é responsável pela aceleração da queda de $S(t)$ ao longo do tempo, de tal maneira que, valores positivos apresentam um decrescimento mais rápido do que valores negativos.

Depois de definir algumas distribuições de probabilidade para descrever o tempo até o retorno do doador, iremos apresentar os procedimentos para o ajuste de modelos de regressão assumindo quaisquer distribuições de probabilidade para esse tempo e relacionando-o com as covariáveis.

Figura 3.4.2. Função de sobrevivência de um modelo paramétrico assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação do doador atribuindo diversos valores para os seus parâmetros.

3.4.1.3 Modelos de Regressão Paramétricos em Análise de Sobrevida

Do ponto de vista paramétrico, os modelos de sobrevida são constituídos por dois componentes: um aleatório e outro determinístico (LOUZADA-NETO; MAZUCHELI; ACHCAR, 2002).

O componente aleatório é dado pela distribuição de probabilidade assumida para o comportamento do tempo de sobrevida. O componente determinístico é dado pelo relacionamento entre os parâmetros desta distribuição e as covariáveis através de uma função positiva e contínua. Genericamente, esta função pode ser escrita por,

$$\eta = g(\alpha \mathbf{x}), \quad (3.30)$$

onde η é um determinado parâmetro de uma distribuição de probabilidade, $g(\cdot)$ é uma função positiva e contínua, geralmente assumida igual a $\exp(\cdot)$, $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k)^T$ é um vetor de coeficientes a serem estimados e associados a um vetor de k covariáveis $\mathbf{x} = (x_1, x_2, \dots, x_k)^T$. Note que $\mathbf{x} = (x_1, x_2, \dots, x_k)^T$ estabelece um efeito multiplicativo no parâmetro η , e este o responsável pela aceleração ou desaceleração do tempo de sobrevida.

Desse modo, a função log-linear é convenientemente utilizada para escrever a relação entre η e o vetor de covariáveis $\mathbf{x}$, de tal maneira que para o i –ésimo doador temos,

$$\ln(\eta(\mathbf{x}_i)) = \alpha_0 + \sum_{j=1}^k \alpha_j x_{ij}, \quad (3.31)$$

A função densidade de probabilidade, descrita no item anterior, é agora denotada por $f(t/\mathbf{x})$, onde seus parâmetros se encontram em função das covariáveis.

A função de sobrevida de um modelo de regressão paramétrico admitindo uma relação do tempo até o retorno do doador, t , com um vetor de k covariáveis relacionadas com o seu perfil, $\mathbf{x}$, através do parâmetro de locação λ , pode ser expressa por,

$$S(t/\mathbf{x}) = P(T > t/\mathbf{x}), \quad (3.32)$$

onde $S(\cdot)$ pode ser estimada a partir de uma determinada distribuição de probabilidade descrita nos itens anteriores informando a probabilidade de um doador retornar para uma doação após o instante t .

Em geral, é comum assumir que as covariáveis afetam apenas o parâmetro de localização de uma determinada distribuição, porém, em muitas aplicações, assumir também que o parâmetro de escala é afetado pelas covariáveis pode ser mais apropriado (MAZUCHELI; LOUZADA-NETO; ACHCAR, 2003).

Conhecida a distribuição de probabilidade e como o ajuste das covariáveis afeta (m) o (s) seu (s) parâmetro (s), seguiremos descrevendo o processo de estimação dos parâmetros por máxima verossimilhança.

3.4.1.4 Estimação de Modelos de Regressão Paramétricos

Seja $t_1, \dots, t_n$ uma amostra aleatória do tempo decorrido entre a primeira e a segunda doação de n doadores de sangue, associado a um vetor de k covariáveis, $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ik})^T$, atribuído ao i -ésimo doador de sangue e seu respectivo vetor de $(k+1)$ coeficientes de regressão, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$, e uma variável indicadora δ_i , onde $\delta_i = 1$ se for observado um retorno do i -ésimo doador de sangue e $\delta_i = 0$ se até o fim do seguimento esse evento não foi observado. A função de verossimilhança, baseada nas informações disponíveis $(t_1, \delta_1, \mathbf{x}_1), (t_2, \delta_2, \mathbf{x}_2), \dots, (t_n, \delta_n, \mathbf{x}_n)$ pode ser escrita como,

$$L(\boldsymbol{\alpha}, \gamma/\mathbf{x}) = \prod_{i=1}^n f(t_i/\mathbf{x}_i)^{\delta_i} S(t_i/\mathbf{x}_i)^{1-\delta_i}, \quad (3.33)$$

onde $f(\cdot)$ é a função densidade de probabilidade e $S(\cdot)$ é a função de sobrevivência apresentadas nos itens anteriores de uma determinada distribuição, α é um vetor de coeficientes associado ao parâmetro de locação através de (3.30), γ é o parâmetro de forma e n é o numero total de doadores.

Observe que a contribuição de cada observação não censurada em (3.33), ou seja, do doador que retornou, é a sua função densidade. Já a contribuição de uma observação censurada em (3.33), ou seja, do doador que não retornou, é a sua função de sobrevivência informando apenas que o tempo para a ocorrência do evento é maior que o período de acompanhamento observado.

Seja $\theta = (\alpha, \gamma)^T$ o vetor de parâmetros associadas a k covariáveis. Os estimadores de máxima verossimilhança de θ podem ser obtidos através da maximização do logaritmo de (3.32) dado por,

$$\begin{aligned} L(\hat{\theta} / \mathbf{x}) &= \max_{\theta} \ln L(\theta / \mathbf{x}) \\ &= \sum_{i=1}^n \left\{ \delta_i \ln [f(t_i / \mathbf{x}_i)] + (1 - \delta_i) \ln [S(t_i / \mathbf{x}_i)] \right\}. \end{aligned} \quad (3.34)$$

A maximização do logaritmo da função de verossimilhança dada (3.34) assumindo as distribuições de probabilidade: Weibull, Log-Normal, Log-Logística e Gama Generalizada Estendida para o tempo decorrido entre a primeira e a segunda doação do doador foram obtidas e se encontram descritas no ANEXO D.

A partir de (3.34), podemos encontrar um sistema de equações não-lineares formado pelas derivadas de cada parâmetro do vetor θ igualado a zero. A solução deste sistema pode ser obtida via métodos numéricos como, por exemplo, o de Newton-Raphson, já implementados nos *softwares* usuais tais como SAS, através do procedimento LIFEREG, ou R, através do procedimento SURVREG.

A estimação da função inversa de $\Sigma = \left[-\partial^2 \ln L(\hat{\theta} / \mathbf{x}) / \partial \theta_i \partial \theta_j \right]$ nos mostrará uma estimativa da matriz de variância-covariância do vetor de parâmetros θ .

Mesmo apresentando uma variedade de distribuições de probabilidade e manejos para manipular a relação das covariáveis e seus parâmetros, em algumas situações práticas tais modelos podem não ser apropriados. Muitas vezes, é comum encontrar uma grande quantidade de indivíduos que após certo período de acompanhamento, podem ser considerados “ímmunes” ao evento de interesse, de tal maneira que a partir dali, o evento é pouco provável ocorrer.

Na área médica, por exemplo, é comum observar num estudo de seguimento de pacientes submetidos a um determinado tratamento, a não ocorrência do evento óbito ou recorrência da doença após certo tempo de acompanhamento. Tais pacientes, livres de sinais ou sintomas da doença, podem ser considerados “curados” e o seu tempo de observação censurado. Não é possível identificar realmente quais são esses indivíduos, apenas inferir a sua presença se o número de censuras nos maiores tempos de acompanhamento for elevado. O interesse principal pode estar justamente aí, em estimar a proporção de “curados”.

Analogamente, no contexto de doação de sangue, é possível pressupor a existência de um número de doadores que após decorrido um determinado tempo da primeira doação, podem ser considerados improváveis de retornar. Neste caso, a construção de modelos que estimam não só a propensão de retorno do doador, mas também a proporção de doadores improváveis de retornar pode ser mais apropriada para descrever o comportamento de retorno do doador, como veremos a seguir.

3.4.1.5 Modelos de Regressão Paramétricos de Longa-Duração em Análise de Sobrevivência

Os modelos de Análise de Sobrevida que incorporam o fato de existirem indivíduos “imunes” ao evento de interesse são conhecidos como modelos de longa-duração ou modelos de mistura (*mixture model*) (MALLER; ZHOU, 1996). Eles possuem a vantagem em relação aos modelos paramétricos usuais por incorporarem a heterogeneidade de duas subpopulações: susceptíveis e imunes ao evento de interesse e com isso prover uma análise aprimorada do efeito das covariáveis estudadas nesses subgrupos separadamente.

Muitas aplicações desses modelos são realizadas em diversas áreas: economia (YAMAGUCHI, 1992), medicina (SPOSTO, 2002), *credit scoring* (ABREU, 2005), *marketing* (QUIDIM, 2005), entre outras.

No contexto de doação de sangue, a construção desse modelo é uma forma de analisar o tempo até o retorno do doador com uma possível ocorrência de “imunidade” ao longo do período de acompanhamento, ou seja, a partir de um determinado tempo é intuitivo supor improvável que os doadores retornem. Tal parcela de doadores não é identificado a priori, mas a presença dos mesmos pode ser inferida se a quantidade de doadores censurados, ou seja, de doadores que não retornaram no estudo, for suficientemente grande a partir de um certo tempo.

Para a definição deste modelo, vamos associar a cada doador uma variável aleatória T representando o tempo até o seu retorno ou até o seu não retorno até o fim do seguimento, com função densidade de probabilidade $f(t)$, onde os parâmetros de locação e de escala são denotados genericamente por λ e γ .

Vamos assumir que os doadores de sangue possam ser classificados em dois grupos: improváveis de retornar (“imunes” ao evento de interesse, ou seja, o retorno) e prováveis de retornar para uma nova doação (“suscetível” ao evento de interesse). Assim como disposto em Maller e Zhou (1996), definimos uma variável aleatória com distribuição

de Bernoulli, B_i , indicando se o i -ésimo doador está em risco de retornar ou não no final do período de seguimento. Dessa maneira, temos que,

$$B_i = \begin{cases} 1, & \text{se o } i\text{-ésimo doador é improvável de retornar} \\ 0, & \text{se o } i\text{-ésimo doador é provável de retornar} \end{cases}, \quad (3.35)$$

onde $i=1,\dots,n$, n é o número de doadores e os valores de B_i são independentes e identicamente distribuídos e ocorrem com probabilidade p se $B_i = 1$ e com probabilidade $1-p$ se $B_i = 0$.

Neste caso, a contribuição na função de verossimilhança agora de um doador que retornou é dada por,

$$P(B_i = 0)f_0(t_i / B_i = 0) = (1-p)f_0(t_i), \quad (3.36)$$

onde $f_0(\cdot)$ é uma determinada função densidade de probabilidade.

De outro modo, também poderíamos partir da probabilidade que o doador tende a retornar até o instante t . Nesse caso, temos que,

$$\begin{aligned} P(T_i < t_i) &= P(B_i = 0)P(T_i < t_i / B_i = 0) + P(B_i = 1)P(T_i < t_i / B_i = 1) \\ &= (1-p)F_0(t_i) + p \times 0, \\ &= (1-p)F_0(t_i) \end{aligned} \quad (3.37)$$

onde $F_0(\cdot)$ é a função de probabilidade acumulada.

Note que derivando (3.37) em relação a t também chegamos em (3.36). Já a contribuição, na respectiva função de verossimilhança, de um doador que está em risco de retornar é dada por,

$$\begin{aligned} P(T_i > t_i) &= P(B_i = 0)P(T_i > t_i / B_i = 0) + P(B_i = 1)P(T_i > t_i / B_i = 1) \\ &= P(B_i = 0)(1 - P(T_i < t_i / B_i = 0)) + P(B_i = 1)P(T_i > t_i / B_i = 1), \\ &= (1-p)S_0(t_i) + p \times 1, \\ &= (1-p)S_0(t_i) + p \end{aligned} \quad (3.38)$$

Os valores de B_i não são observados, mas a presença de doadores improváveis de retornar pode ser inferida se a quantidade de doadores que não retornaram para sua segunda doação (censurados) no final do estudo for suficientemente grande.

Reescrevendo (3.38), temos que a função de sobrevivência de um modelo de longa-duração será expressa por,

$$S(t/\mathbf{x}) = p + (1-p)S_0(t/\mathbf{x}), \quad (3.39)$$

onde $S(\cdot)$ é a função de sobrevivência global para os doadores, $S_0(\cdot)$ é a função de sobrevivência dada em (3.32) e $0 < p < 1$ representa a proporção de doadores improváveis de retornar. Se $S_0(0/\mathbf{x}) = 1$, temos $S(0/\mathbf{x}) = 1$, e, considerando $S_0(\infty/\mathbf{x}) = 0$, temos $S(\infty/\mathbf{x}) = p$, ou seja, à medida que as observações de t assumem valores grandes, a função de sobrevivência S aproxima-se de p .

Observe que (3.39) pode ser considerado um modelo de mistura com dois componentes associados a duas subpopulações de doadores distintas: o primeiro com proporção p e o segundo associado ao componente paramétrico, $S_0(\cdot)$. Note também que se assumirmos $p = 0$, estaremos considerando o modelo paramétrico usual, onde $S(t/\mathbf{x}) = S_0(t/\mathbf{x})$, ou seja, descartamos a estimação de uma proporção de indivíduos improváveis de retornar, dado em (3.32).

Por sua vez, a função densidade de probabilidade pode ser reescrita de (3.36) como,

$$f(t/\mathbf{x}) = (1-p)f_0(t/\mathbf{x}). \quad (3.40)$$

onde $f_0(\cdot)$ é uma determinada função densidade de probabilidade.

Devido o fato de trabalhar com duas subpopulações distintas, é esperado que algumas características estejam mais relacionadas com um determinado subgrupo do que

outro, de tal maneira que, o perfil de um doador que retornou para uma segunda doação e um doador que não retornou não necessariamente são os mesmos. Ao assumirmos que o efeito das covariáveis, associadas com o perfil do doador, afeta ambos os componentes da equação (3.39), estamos dizendo que a proporção de doadores improváveis de retornar, p , e a distribuição dos tempos até o próximo retorno do doador podem ser diferenciadas de acordo com o seu perfil.

Nesse caso, a função de sobrevivência do modelo de longa-duração pode ser reescrito de (3.39) da seguinte forma,

$$S(t/\mathbf{x}, \mathbf{z}) = p(\mathbf{z}) + (1 - p(\mathbf{z}))S_0(t/\mathbf{x}), \quad (3.41)$$

onde $S_0(\cdot)$ é a função de sobrevivência descrita em (3.32) e $0 < p(\mathbf{z}) < 1$, como já mencionado, representa a proporção de doadores improváveis de retornar, a qual é também dependente de um vetor de m covariáveis, $\mathbf{z} = (z_1, z_2, \dots, z_m)^T$, associado a um vetor de $m+1$ parâmetros, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_m)^T$, podendo $\mathbf{x}$ e $\mathbf{z}$ serem iguais.

A respectiva função densidade de probabilidade, reescrita de (3.40), é dada por,

$$f(t/\mathbf{x}, \mathbf{z}) = (1 - p(\mathbf{z}))f_0(t/\mathbf{x}). \quad (3.42)$$

Podemos modelar a dependência da proporção de doadores improváveis de retornar, p , e do vetor de m covariáveis, $\mathbf{z}$, através de um componente logístico, assim como disposto em Peng, Dear e Denham (1998), tal que,

$$1 - p(\mathbf{z}) = \exp(\mathbf{z}\boldsymbol{\beta}) / (1 + \exp(\mathbf{z}\boldsymbol{\beta})), \quad (3.43)$$

onde $\boldsymbol{\beta}$ é o vetor de $m+1$ parâmetros.

O comportamento da função de sobrevivência de um modelo longa-duração, dado em (3.39), assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação do doador, dado em (3.29), pode ser visto na Figura 3.4.3..

Vários casos foram simulados: 1) valores α , γ e p fixados em 2,5, -1 e 0,3 respectivamente e atribuídos diferentes valores para o parâmetro de locação λ (Figura 3.4.3. – painel superior esquerdo), 2) valores de α e γ fixados em 2,5 e -1 respectivamente e atribuídos diferentes valores para a proporção, p , e para o parâmetro de locação λ , porém próximos dentre eles (Figura 3.4.3. – painel superior do meio), 3) situação semelhante ao caso dois, porém assumindo interações (Figura 3.4.3. – painel superior direito), 4) valores de λ , α e γ fixados em 2,5, 2,5 e -1 respectivamente e atribuídos diferentes valores para a proporção, p (Figura 3.4.3. – painel inferior esquerdo), 5) valores de α e γ fixados em 2,5 e -1 respectivamente e atribuídos diferentes valores para a proporção, p , e para o parâmetro de locação λ (Figura 3.4.3. – painel inferior do meio) e 6) situação semelhante ao caso cinco, porém assumindo interações (Figura 3.4.3. – painel inferior direito).

Nos casos (1), (2) e (3), à medida que as observações de t assumem valores grandes, a função de sobrevivência $S(t)$ aproxima-se de p . No contexto de doação de sangue, o primeiro caso corresponderia a uma situação onde o perfil do doador é determinante na distinção do seu retorno, de tal maneira que esse evento tende a ser freqüente (determinado pela queda de $S(t)$) até um determinado valor de t , em contrapartida, é esperado o mesmo número de doadores que jamais retornarão. Já os casos (2) e (3) correspondem a situações onde, independente do perfil do doador, o comportamento de retorno é aproximadamente o mesmo.

Nos casos (4), (5) e (6), diferentes valores de λ caracterizam uma queda acentuada na função de sobrevivência $S(t)$ em distintos valores de t até convergir para a proporção p , caracterizando o perfil dos doadores improváveis de retornar. No contexto de doação de sangue, o caso (4) corresponderia a uma situação onde entre os doadores que

Figura 3.4.3. Função de sobrevivência de um modelo paramétrico longa-duração assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação do doador atribuindo diversos valores para os seus parâmetros.

retornaram, o fizeram até o tempo t , independente do seu perfil (covariáveis), porém, a proporção de doadores improváveis de retornar depende do perfil do candidato. Já nos casos (5) e (6) correspondem a situações onde tanto o comportamento de retorno do doador quanto o número esperado de doadores improváveis de retornar depende do seu perfil.

A estimação dos parâmetros de um modelo longa-duração pode ser feita através do método de máxima verossimilhança conforme descrito a seguir.

3.4.1.6 Estimação de Modelos de Regressão de Longa-Duração Paramétricos

Seja $t_1, \dots, t_n$ uma amostra aleatória do tempo decorrido entre a primeira e a segunda doação de n doadores de sangue, associado ao vetor de k covariáveis $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{ki})^T$ e m covariáveis $\mathbf{z}_i = (z_{1i}, z_{2i}, \dots, z_{mi})^T$ atribuído ao i -ésimo doador, e uma variável indicadora δ_i , onde $\delta_i = 1$ se for observado um retorno do i -ésimo doador de sangue na amostra e $\delta_i = 0$ se não for observado esse evento. A função de verossimilhança, baseada nas informações disponíveis $(t_1, \delta_1, \mathbf{x}_1, \mathbf{z}_1), (t_2, \delta_2, \mathbf{x}_2, \mathbf{z}_2), \dots, (t_n, \delta_n, \mathbf{x}_n, \mathbf{z}_n)$ pode ser escrita como,

$$L(\alpha, \beta, \gamma / \mathbf{x}, \mathbf{z}) = \prod_{i=1}^n f(t_i / \mathbf{x}_i, \mathbf{z}_i)^{\delta_i} S(t_i / \mathbf{x}_i, \mathbf{z}_i)^{1-\delta_i}, \quad (3.44)$$

onde $f(\cdot)$ é a função densidade de probabilidade definida em (3.42), $S(\cdot)$ é a função de sobrevivência definida em (3.41) e n é o numero total de doadores.

Seja $\theta = (\alpha, \beta, \gamma)^T$ o vetor de parâmetros associadas a k e m covariáveis respectivamente para α e β . Os estimadores de máxima verossimilhança de θ podem ser obtidos através da maximização do logaritmo de (3.44), ou seja,

$$L(\hat{\alpha}, \hat{\beta}, \hat{\gamma} / \mathbf{x}, \mathbf{z}) = \max_{\alpha, \beta, \gamma} \ln L(\alpha, \beta, \gamma / \mathbf{x}, \mathbf{z}) \quad (3.45)$$

$$= \sum_{i=1}^n \left\{ \delta_i \ln [f(t_i / \mathbf{x}_i, \mathbf{z}_i)] + (1 - \delta_i) \ln [S(t_i / \mathbf{x}_i, \mathbf{z}_i)] \right\}$$

A maximização do logaritmo da função de verossimilhança dada (3.45) assumindo as distribuições de probabilidade: Weibull, Log-Normal, Log-Logística e Gama Generalizada Estendida para o tempo decorrido entre a primeira e a segunda doação do doador foram obtidas e se encontram descritas no ANEXO E.

A partir de (3.45), podemos encontrar um sistema de equações não-lineares formado pelas derivadas de cada parâmetro do vetor θ igualado a zero. Porém, pode ser custoso obter a solução desse sistema diretamente por métodos iterativos. Uma alternativa é maximizar (3.45) encontrando pontos de máximo locais por algoritmos de otimização como sugerido por Mazucheli, Louzada-Neto e Achcar (2003). Tal método pode ser implementado no *software* SAS através do procedimento NLP. Outra sugestão para o ajuste de modelos com covariáveis, é a biblioteca *gfcure*, desenvolvido por Peng em (PENG; DEAR; DENHAM, 1998) com versões para os programas S-Plus e R.

A estimação da função inversa de $\Sigma = \left[-\partial \ln L(\hat{\theta} / \mathbf{x}) / \partial \theta_i \theta_j \right]$ nos mostrará uma estimativa da matriz de variância-covariância do vetor de parâmetros θ .

3.4.1.7 Medidas de Avaliação de Modelos em Análise de Sobrevida

A forma mais simples e eficiente de selecionar o modelo que “melhor” se ajusta ao conjunto de dados é através de técnicas gráficas. (MALLER; ZHOU, 1996; COLOSIMO; GIOLO, 2006). Uma delas consiste em obter a função de sobrevivência empírica a partir do estimador de Kaplan-Meier, denotada aqui por $\hat{S}_{KM}(t)$, e em seguida,

estimar a mesma função assumindo uma determinada distribuição de probabilidade, por exemplo, uma distribuição Weibull denotada aqui por $\hat{S}_w(t)$.

Posteriormente, comparam-se graficamente as funções de sobrevivência estimadas para cada distribuição de probabilidade com $\hat{S}_{KM}(t)$. O “melhor” modelo ajustado será aquele cujos pontos no gráfico estiverem mais próximos da reta $y = x$, onde $x = \hat{S}_{KM}(t)$ e $y = \hat{S}_w(t)$, por exemplo.

A utilização do critério AIC (*Akaike Information Criterion*) também é um meio de seleção de modelos. Valores pequenos do critério AIC indicam melhor qualidade do ajuste entre os comparados, podendo ser definido como,

$$AIC = -2 \times \left[\max(\ln L(.)) \right] + 2c. \quad (3.46)$$

onde $\max(\ln L(.))$ é dado pela maximização do logaritmo da função de verossimilhança dada em (3.34) ou (3.45) e c é o número de parâmetros do modelo em questão.

Este critério é bastante utilizado em problemas de estatística aplicada, no entanto, ele tende a superestimar o efeito do número de parâmetros quando o tamanho amostral é pequeno, ou em situações onde o número de parâmetros é grande em relação ao tamanho amostral. Nestes casos, o AIC com correção proposta por Sugiura (1978) é uma alternativa viável.

Por outro lado, quando o tamanho amostral é grande, a contribuição relativa do “termo de penalidade” associada ao número de parâmetros ($2c$) na determinação do AIC torna-se pequena quando comparada ao valor de $(\ln L(.))$. Neste caso, uma alternativa é o *Bayesian Information Criterion* (BIC), dado por,

$$BIC = -2 \times \left[\max(\ln L(.)) \right] + c \ln(n). \quad (3.47)$$

que inclui o tamanho amostral n em seu “termo de penalidade” (SCHWARZ, 1978).

CAPÍTULO IV

Resultados

4.1 Descrição do Histórico de Doação

O histórico de doação registrado em banco de dados de toda a região de abrangência de Ribeirão Preto (Ver Anexo A) é composto por 928173 registros de doação. Destas, 21% não chegaram a ser efetivamente uma doação: cerca de 20% eram referentes a um episódio do doador inapto temporariamente, 0,9% eram referentes a um episódio do doador inapto definitivamente e 0,005% não tinham a condição de aprovado na triagem.

De acordo com este histórico, nota-se um aumento considerável no número de doações de 1996 ao final de 2002, estabilizando do início de 2002 ao final de 2004, com tendência a declinar a partir deste ponto (Figura 4.1.1.). Porém, quando associamos o número de intenções de doação com o tipo de doação (Figura 4.1.2.), verifica-se que esse aumento é característico da doação voluntária, enquanto que as doações autólogas, específicas e vinculadas apresentam-se em geral um comportamento constante ao longo do tempo, porém, todas com a existência de alguns picos expressivos.

Figura 4.1.1. Número de doações mensais realizadas no período de 1996 a 2005 obtidas na região de abrangência de Ribeirão Preto.

Figura 4.1.2. Número de doações mensais realizadas no período de 1996 a 2005 obtidas na região de abrangência de Ribeirão Preto de acordo com o tipo de doação: autóloga (1°), específica (2°), vinculada e voluntária (3°).

A ocorrência dos maiores picos (Figura 4.1.1) é obtida geralmente no mês de novembro (anos de 1997, 1998, 1999, 2002, 2003), quando tradicionalmente ocorrem as campanhas de motivação, realizadas durante a semana do Doador de Sangue (3ª semana do mês). Já os meses de fevereiro e junho caracterizam-se geralmente pelo menor número de doações. Isto acaba sendo um problema enfrentado pelos centros de hemoterapia, pois esses meses correspondem ao período de férias e feriados, onde geralmente há uma demanda maior por bolsas de sangue coincidindo com um aumento considerável no tráfego nas rodovias e um conseqüente aumento no número de acidentes com vítimas que necessitam de transfusões.

O número de doadores registrados em banco de dados em toda a região de abrangência de Ribeirão Preto durante o período estudado é de 367869. Destes, 30% não foram aprovados na doação: cerca de 28% deles foram considerados inaptos temporariamente, 2% deles foram considerados inaptos definitivamente e 0,01% não tinham a condição de aprovado na triagem. A Figura 4.1.3. apresenta o número de doadores de acordo com o mês da sua primeira doação. Através dela podemos notar que este número se mantém constante ao longo do período estudado, apresentando um pequeno decréscimo a partir de 2004.

Figura 4.1.3. Número de doadores que iniciaram seu ciclo de doação mensalmente no período de 1996 a 2005 obtidas na região de abrangência de Ribeirão Preto.

Apesar da captação de novos doadores ao longo do tempo ser constante, quando associamos o número de doadores com o tipo de doação, verificamos um aumento a partir de 2001 de doadores voluntários com relação ao número de doadores vinculados. Doadores autólogos e específicos se mantiveram constantes em todo o período estudado (Figura 4.1.4.).

Figura 4.1.4. Número de doadores mensais registrados no período de 1996 a 2005 da região de abrangência de Ribeirão Preto de acordo com o tipo de doação: autóloga (1°), específica (2°), vinculada e voluntária (3°).

4.2 Descrição do Histórico de Doação dos Doadores Voluntários

Para estudar o comportamento de retorno dos doadores de sangue que manifestaram sua primeira intenção de doar em caráter voluntário, através do tempo entre a primeira e a segunda doação, foram excluídos os doadores que retornaram antes do término do intervalo mínimo entre as doações¹³, dado que não é esperado que algum doador retorne antes do final deste período estipulado em no mínimo 60 dias. Assim, obteve-se uma casuística final de 115553 doadores, dos quais 57,1% retornaram e iniciaram o segundo ciclo de doação. O histórico de doação desse grupo de doadores é composto por 346211 doações, com uma taxa retorno/doador de 1,996. O máximo de retornos realizados por um doador foi 39.

A Figura 4.2.1. descreve o perfil de retorno de cada doador voluntário que retornou pelo menos uma vez em algum serviço de hemoterapia localizado na região estudada, de tal maneira que podemos identificar o número de retornos realizados e o tempo observado das suas ocorrências acumulado. Note que a amplitude de variação do tempo necessário para a realização de um determinado número de retornos é grande. O tempo de retorno para a segunda doação, por exemplo, varia de 61 a 3084 dias. Isso sugere que os doadores de sangue tendem a apresentar um comportamento de retorno extremamente variável.

¹³ Apesar de não ser esperado que um doador retorne antes do intervalo mínimo entre as doações, é possível que isso ocorra como visto alguns casos no banco de dados. Porém, tais casos sugerem muitas vezes situações particulares, que deveriam ter sido profundamente averiguadas. Por exemplo, doadores impossibilitados de doar dado o período entre as intenções extremamente curto, doadores dos quais realmente foram coletadas bolsas de sangue em um curto espaço de tempo, doadores convocados para fazer exames antes de uma doação por aférese, entre outros. Como a coleta de dados para este estudo foi feita retrospectivamente, não pudemos justificar os retornos observados antes do término desse período.

Figura 4.2.1. Número de retornos realizados no período de 1996 a 2005 de acordo com o tempo de acompanhamento dos doadores voluntários.

4.3 Descrição do Número de Doadores Voluntários de acordo com as variáveis sócio-demográficas.

Entre os doadores de sangue voluntários estudados, cerca de 50,3%, realizaram sua primeira doação por meio da coleta externa,. A captação interna foi responsável por 49,7% dos doadores, sendo subdivididos em 6,3% obtidos por meio das Unidades Pequenas e 43,3% obtidos por meio das Unidades Grandes.

A Tabela 4.3.1 sumariza a distribuição de algumas variáveis relacionadas com o perfil do doador. Entre os doadores, 60,9% eram do sexo masculino e 39,1% eram do sexo feminino, sendo distribuídos nas seguintes faixas etárias: 15,7% entre 18 a 19 anos, 38,1% entre 20 a 29 anos, 24,5% entre 30 a 39 anos, 15,3% entre 40 a 49 anos, 6,1% entre 50 a 59 anos e 0,3% entre 60 a 65 anos. Quanto à escolaridade, foi constatado que: 0,4% eram analfabetos, 27,2% possuíam 1º grau incompleto, 17,2% possuíam 1º grau completo ou 2º grau incompleto, 46,0% possuíam 2º grau completo ou 3º grau incompleto, 8,8% possuíam 3º grau completo e 0,4% possuíam pós graduação. Também foi observado que 0,6% foram declarados da cor amarela, 76,2% foram declarados da cor branca e 23,2% foram declarados da cor negra. No que se refere ao estado civil, 44,8% eram casados, 4,8% eram divorciados, 47,3% eram solteiros, 1,2% eram viúvos e 1,9% foram categorizados na opção outros .

As frequências dos tipos sanguíneos na casuística foi 37,6%, 3,7% 11,2% e 47,4%, respectivamente, para doadores com sangue tipo A, AB, B e O, sendo que 87,3% destes apresentam fator Rh positivo (Tabela 4.3.1).

Do total de doadores estudados, 57,1% retornaram e iniciaram o segundo ciclo de doação. Note que a distribuição das variáveis sócio-demográficas desse subgrupo de doadores é aproximadamente a mesma do subgrupo de doadores que não retornaram,

Tabela 4.3.1. Descrição das variáveis sócio-demográficas dos doadores voluntários do Centro Regional de Hemoterapia de Ribeirão Preto.

Covariáveis	Retorno				Total	
	Não		Sim			
	n ₁ (49594)	% (42,9)	n ₂ (65959)	% (57,1)	N (115553)	% (100,0)
Sexo						
Feminino	20034	40,4	25109	38,1	45143	39,1
Masculino	29560	59,6	40850	61,9	70410	60,9
Cor da Pele ^(a)						
Amarela	346	0,7	306	0,5	652	0,6
Branca	37665	76,6	49946	75,9	87611	76,2
Negra	11145	22,7	15515	23,6	26660	23,2
Faixa Etária ^{(a) (b)}						
18 a 19	7658	15,4	10457	15,9	18115	15,7
20 a 29	19941	40,2	24049	36,5	43990	38,1
30 a 39	11492	23,2	16783	25,4	28275	24,5
40 a 49	7245	14,6	10389	15,7	17634	15,3
50 a 59	2974	6,0	4114	6,2	7088	6,1
60 a 65	259	0,5	149	0,2	408	0,3
Escolaridade ^(a)						
Analfabeto	131	0,4	240	0,5	371	0,4
1º Grau Incompleto	8420	24,3	15214	29,1	23634	27,2
1º Grau Completo/2º Grau Incompleto	5870	17,0	9033	17,3	14903	17,2
2º Grau Completo/3º Grau Incompleto	17193	49,7	22727	43,5	39920	46,0
3º Grau Completo	2836	8,2	4846	9,3	7682	8,8
Pós Graduação	122	0,3	206	0,4	328	0,4
Estado Civil ^(a)						
Casado	20166	41,2	31157	47,5	51323	44,8
Divorciado	2066	4,2	3458	5,3	5524	4,8
Outros	821	1,7	1331	2,0	2152	1,9
Solteiro	25387	51,8	28801	43,9	54187	47,3
Viúvo	548	1,1	810	1,2	1358	1,2
Tipo Sanguíneo ^(a)						
A	18824	38,3	24475	37,1	43299	37,6
AB	1736	3,5	2517	3,8	4253	3,7
B	5481	11,2	7428	11,3	12909	11,2
O	23091	47,0	31458	47,7	54549	47,4
Fator Rh ^(a)						
Negativo	5583	11,4	8976	13,6	14559	12,7
Positivo	43519	88,6	56883	86,4	100402	87,3

(a) Variáveis que contêm dados faltantes.

(b) Variáveis obtidas na primeira doação.

exceto pelo ligeiro aumento do número de indivíduos com 1º grau incompleto paralelamente com uma diminuição do número de indivíduos com 2º grau completo ou 3º grau incompleto assim como um aumento do número de casados e uma diminuição do número de solteiros entre os doadores que retornaram (Tabela 4.3.1.).

Entre os doadores que realizaram sua primeira doação por meio da coleta externa, 57,5% retornaram, da mesma forma que 60,7% e 56,0% retornaram entre aqueles que realizaram sua doação por meio da coleta interna, respectivamente em uma das unidades pequenas e unidades grandes.

Com o intuito de descrever as associações entre as categorias destas variáveis relacionadas com o perfil do doador de sangue, utilizamos uma análise de correspondência múltipla¹⁴ (GREENACRE, 1993). O agrupamento dessas variáveis realizado pela análise de correspondência, evidencia principalmente através da associação entre estado civil, faixa etária e nível de escolaridade, alguns perfis de doador¹⁵ (ANEXO F).

Nesse caso, doadores declarados da cor amarela se encontram mais associados com o estado civil solteiros, com até 30 anos de idade e um nível de escolaridade entre 2º grau completo/3º grau incompleto ou pós-graduação.

Outro fato que evidencia é o fenômeno natural do aumento da faixa etária e a mudança do *status* do doador com relação ao estado civil, alterando de casado para divorciado à viúvo. Também com ele se observa a diminuição do nível de escolaridade à medida que aumenta a faixa etária.

¹⁴ A análise de correspondência é uma técnica estatística multivariada que permite medir e visualizar, de modo gráfico, o grau de associação entre um conjunto de categorias de variáveis qualitativas. (GREENACRE, 1993).

¹⁵ A inclusão da variável nível de escolaridade, apesar da grande quantidade de *missings*, não alterou drasticamente as conclusões da análise sem ela, vindo a identificar melhor o perfil do doador de sangue, apesar da perda de informação.

O fato do doador se submeter a uma segunda doação não lhe atribui um perfil diferenciado, de tal maneira que, tanto as categorias mais associadas ao retorno do doador (representados a direita do gráfico) quanto aquelas mais associadas ao não retorno do doador (representadas a esquerda do gráfico) estão dispostas próximas da origem (valor zero).

Isso complementa o que foi dito anteriormente, evidenciando mais uma vez que a distribuição das variáveis sócio-demográficas entre os doadores que retornaram ou não para uma subsequente doação é aproximadamente a mesma.

A representatividade dos tipos sanguíneos de acordo com o fator Rh entre os doadores voluntários da região de abrangência de Ribeirão Preto é menor em alguns casos, principalmente entre os Rh negativos quando comparado com o esperado na população brasileira (Tabela 4.3.2.).

Dessa maneira, a captação dos tipos sanguíneos e respectivos fatores Rh: A-, A+, AB- e O- é abaixo do esperado encontrar na população e dos tipos AB+, B+ e O+ é acima do esperado encontrar na população.

Tabela 4.3.2. Distribuição do tipo sanguíneo e fator Rh dos doadores voluntários do Centro Regional de Hemoterapia de Ribeirão Preto e do esperado encontrar na população.

Tipo Sanguíneo/Fator Rh	Casuística (%)	População (%)
A-	4,7	6% a 8%
A+	32,9	34% a 36%
AB-	0,4	0,5% a 1%
AB+	3,3	2,5% a 3%
B-	1,3	1 a 2%
B+	9,9	8 a 9%
O-	6,2	7 a 9%
O+	41,2	36 a 38%

Fonte: Hospital da Aeronáutica dos Afonsos.

Nas Tabelas 4.3.3. a 4.3.5. estão um sumário das variáveis relacionadas com o comportamento e motivação do doador associado com a sua segunda doação.

Entre os doadores estudados, a maioria tende a retornar pela segunda vez em caráter voluntário, cerca de 80,0% (Tabela 4.3.3.).

Tabela 4.3.3. Distribuição dos doadores voluntários que retornaram para sua segunda doação de acordo com o tipo de doação.

Tabela 4.3.3.

1ª Doação	2ª Doação n (%)			Total
	Específica	Vinculada	Voluntária	
Voluntária	10 (0,02)	10921 (20,1)	43334 (79,9)	54265*

* Total de doadores de sangue parcial dado que indivíduos que foram reprovados temporariamente ou definitivamente na segunda intenção de doação não respondem essa questão. 11 doadores aprovados nas duas condições não tinham o tipo de doação na segunda intenção.

Também observa-se que a porcentagem de aprovação na triagem na segunda doação foi cerca de 82%, o que corresponde a aproximadamente 47% do número total de doadores voluntários estudados (Tabela 4.3.4.).

Tabela 4.3.4. Distribuição dos doadores voluntários que retornaram para sua segunda doação de acordo com a condição de triagem.

1ª Doação	2ª Doação n (%)			Total
	AP	RT	RD	
AP	54276 (82,3)	11367 (17,2)	316 (0,5)	65959 (100,0)

Doadores de sangue voluntários tendem a retornar no mesmo tipo de unidade de coleta que doaram na primeira vez. Isso foi observado entre 83,3% dos doadores que iniciaram seu ciclo de doação na coleta externa, 96,5% entre aqueles que iniciaram seu ciclo em uma das Unidades Pequenas e 92,2% entre aqueles que iniciaram seu ciclo em uma das Unidades Grandes (Tabela 4.3.5.).

Tabela 4.3.5. Distribuição dos doadores voluntários que retornaram para sua segunda doação de acordo com a unidade.

1ª Doação	2ª Doação n (%)			Total
	Coleta Externa	Unidades Pequenas	Unidades Grandes	
Coleta Externa	27852 (83,3)	178 (0,5)	5422 (16,2)	33452 (50,7)
Unidades Pequenas	56 (1,3)	4281 (96,5)	98 (2,2)	4435 (6,7)
Unidades Grandes	2077 (6,9)	112 (2,4)	25883 (92,2)	28072 (42,6)
Total	29985 (45,5)	4571 (6,9)	31403 (47,6)	65959 (100,0)

Somando-se a isso, podemos observar que o tempo entre a primeira e a segunda doação é inversamente relacionado com o número de doações realizadas ao longo do seu histórico de doação (Tabela 4.3.6.). Assim, doadores que apresentam um curto espaço de tempo entre as duas primeiras doações, são aqueles que tendem a retornar um numero maior de vezes e vice versa (Figura 4.3.1.). Entre os doadores que retornaram, 33,7% o fizeram apenas uma vez, 41,8% o fizeram até quatro vezes, 16,3% de cinco a oito vezes, 5,2% de nove a doze vezes e 2,9% mais que doze vezes.

Tabela 4.3.6. Descrição do tempo entre a primeira e a segunda doação do doador voluntário de acordo com o número de retornos realizados ao longo do histórico de doação.

	Número de Retornos									
	1		2 a 4		5 a 8		9 a 12		>12	
	n (22240)	% (33,7)	n (27599)	% (41,8)	n (10778)	% (16,3)	n (3438)	% (5,2)	n (1904)	% (2,9)
<i>Tempo</i> ^(a)										
Até 180 dias	6166	27,72	10951	39,68	5422	50,31	2095	60,94	1358	71,32
Até 360 dias	6056	27,23	8289	30,03	3177	29,48	925	26,91	404	21,22
Até 720 dias	5409	24,32	5441	19,71	1605	14,89	344	10,01	125	6,57
Mais que 720 dias	4609	20,72	2918	10,57	574	5,33	74	2,15	17	0,89

(a) Tempo entre a primeira e a segunda doação.

Figura 4.3.1. Porcentagem de doadores voluntários de acordo com o número total de retornos realizados e o tempo entre a primeira e a segunda doação.

Na Tabela 4.3.7. se encontra a distribuição do tempo até o retorno do doador ou tempo de acompanhamento até o fim do seguimento (Figura 4.3.2). Entre os doadores que retornaram, cerca de 88% o fizeram em até dois anos após a primeira doação.

Tabela 4.3.7. Distribuição do tempo entre a primeira e a segunda doação ou até o fim do seguimento do doador voluntário.

Tempo	Retorno					
	Não			Sim		
	n (49594)	%	% Acumulada	n (65959)	%	% Acumulada
180	3439	6,9	6,9	25992	39,4	39,4
360	5084	10,2	17,1	18851	28,6	68,0
720	8759	17,7	34,8	12924	19,6	87,6
1080	7610	15,3	50,1	4402	6,7	94,3
1440	6836	13,8	63,9	1984	3,0	97,3
1800	5097	10,3	74,2	949	1,4	98,7
2160	3986	8,0	82,2	496	0,8	99,5
2520	2952	6,0	88,2	242	0,4	99,8
2880	3196	6,4	94,6	96	0,2	100,0
3240	2596	5,2	99,8	23	0,0	100,0
3300	39	0,1	99,9	0	0,0	100,0

Figura 4.3.2. Porcentagem acumulada do tempo entre a primeira e a segunda doação do doador voluntário ou até o fim do seguimento o doador voluntário.

4.4. Descrição do tempo entre a primeira e a segunda doação dos Doadores Voluntários

Para a descrição do tempo entre a primeira e a segunda doação do doador voluntário e a obtenção da sua propensão à não retornar (taxa de não retorno)¹⁶ foram construídas curvas a partir do estimador de Kaplan-Meier. Observe que essa estimativa permanece igual a um durante o intervalo mínimo entre as doações e começa a decair dali por diante, destacando uma queda acentuada até 720 dias após a sua primeira doação tendendo a estabilizar posteriormente (Figura 4.4.1.).

Após um ano do início do ciclo de doação, 59,8% dos doadores não haviam retornado, decaindo para aproximadamente 46,5% após dois anos de acompanhamento (acréscimo de aproximadamente 13,3% de doadores que retornaram). A Figura 4.4.1 evidencia ainda que após três anos decorridos da primeira doação, 41,0% não retornaram, bem como 37,8% após quatro anos e 35,7% após cinco anos de acompanhamento. A partir do sexto ano, a função de sobrevivência perde a sua inclinação indicando que raramente um doador retorna para uma segunda doação, dado que não retornou seis anos após a primeira doação. No final do seguimento, 31,8% dos doadores não haviam retornado (acréscimo de aproximadamente 14,7% de doadores que retornaram entre dois à nove anos após a primeira doação).

Uma aproximação da função risco foi obtida a partir do método atuarial, admitindo intervalos de amplitude de 30 dias (Figura 4.4.2.). A função risco representa a chance instantânea do doador retornar em todo o período.

¹⁶ Denominaremos por taxa de não retorno o valor da propensão à não retornar (estimado pelo método de Kaplan-Meier) vezes 100.

Figura 4.4.1. Função de sobrevivência representando a propensão à não retornar dos doadores voluntários.

Figura 4.4.2. Função Risco instantâneo à retornar. Os picos mais expressivos de retorno correspondem aproximadamente a 4, 12 e 24 meses após o início do ciclo de doação.

A partir dela, podemos observar que um doador atinge a maior chance de retornar no período de 90 a 120 dias após sua primeira doação, decrescendo exponencialmente até o final do período de seguimento, mas com a presença de picos expressivos aproximadamente um ano e dois anos após a primeira doação.

A Figura 4.4.3. mostra as estimativas de Kaplan-Meier segundo unidade de coleta, sendo seus resultados sumarizados na Tabela 4.4.1.. A Figura 4.4.4. apresenta uma aproximação da perspectiva função risco pelo método atuarial.

A taxa de não retorno entre os doadores que realizaram sua primeira doação em uma das unidades pequenas foi a menor em todo o período estudado. No final do seguimento, 27,9% e 32,1% dos doadores que realizaram sua primeira doação em uma das unidades pequenas e grandes e 32,3% dos doadores que realizaram sua primeira doação por meio da Coleta Externa não haviam retornado.

Independente da unidade de coleta, a chance do doador de sangue retornar é maior no período de 90 a 120 dias após o início do ciclo de doação, principalmente dos doadores advindos das unidades pequenas (Figura 4.4.4.). Após este período a chance do doador retornar é aproximadamente a mesma.

A Figura 4.4.5. mostra as estimativas de Kaplan-Meier estratificadas por cada uma das covariáveis relacionadas com o perfil do doador, sendo seus resultados sumarizados na Tabela 4.4.1..

Dentre o conjunto de doadores estudados, 55,6% do sexo feminino e 58,0% do sexo masculino retornaram para uma segunda doação. Após um ano decorrido da primeira doação, 60,1% dos doadores do sexo feminino e 59,6% dos doadores do sexo masculino não haviam retornado. De maneira geral, o retorno dos doadores de ambos os sexos são muito próximos principalmente nos primeiros anos de seguimento. No final do seguimento, 33,5% e 30,9% dos doadores do sexo feminino e masculino não haviam retornado.

Figura 4.4.3. Função de sobrevivência representando a propensão à não retornar dos doadores voluntários segundo unidade de coleta.

Figura 4.4.4. Função Risco instantâneo à retornar segundo unidade de coleta. Os picos mais expressivos de retorno correspondem aproximadamente a 4, 12 e 24 meses após o início do ciclo de doação.

Entre os indivíduos declarados como amarelos, brancos e negros, a porcentagem de retorno foi de aproximadamente 46,9%, 57,0% e 58,2% respectivamente. Em todo o período estudado, a taxa de não retorno para indivíduos declarados negros ou brancos foi menor quando comparado com os indivíduos declarados amarelos. No fim do seguimento, 42,1%, 32,2% e 29,6% dos doadores declarados como amarelos, brancos e negros respectivamente não haviam retornado para a segunda doação.

Com exceção da faixa etária de sessenta anos a sessenta e cinco anos, as demais faixas apresentaram um número de retornos acima de 54%, destacando a faixa dos 30 aos 39 anos com o maior número de retornos, cerca de 59,4%. A taxa de não retorno da faixa etária de sessenta anos ou mais é a maior em todo o período estudado, já as demais faixas etárias apresentam um comportamento aproximadamente similar em todo o período estudado.

Dentre as categorias correspondentes ao estado civil, a opção divorciado apresentou um maior número de retornos, cerca de 62,6%, enquanto que a opção solteiros apresentou o menor número, cerca de 53,2%. Exceto a categoria solteiros, o comportamento de retorno dos demais doadores são muito próximos em todo o período estudado, destacando porém a opção viúvo que depois de certo tempo, tende a apresentar um comportamento mais parecido com o de solteiros. No fim do seguimento, o número de doadores solteiros e viúvos que não retornaram são próximos (36,4% e 33,0% respectivamente), bem como o número de doadores casados, divorciados e pertencentes a categoria outros (27,7%, 27,3% e 25,0% respectivamente).

O número de indivíduos que retornaram para uma segunda doação com tipo sanguíneo A, AB, B e O foram respectivamente 56,5%, 59,2%, 57,5% e 57,7% e o comportamento de retorno destes também são muito semelhantes em todo o período de acompanhamento.

Tabela 4.4.1. Sumário das estimativas de Kaplan-Meier para cada uma das variáveis sócio-demográficas dos doadores voluntários.

Covariáveis	n ^(c)	Retorno ^(d)	Taxa de Não Retorno (%)			
			Dias			
			180	360	720	3240
	115553 (100,0)	65959 (57,1%)	77,1	59,8	46,5	31,8
<i>Sexo</i>						
Feminino	45143 (39,1)	25109 (55,6)	78,1	60,1	47,1	33,5
Masculino	70409 (60,9)	40850 (58,0)	76,5	59,6	46,2	30,9
<i>Cor da Pele</i> ^(a)						
Amarela	652 (0,6)	306 (46,9)	85,0	71,4	60,1	42,1
Branca	87610 (76,2)	49946 (57,0)	77,5	60,2	46,8	32,2
Negra	26660 (23,2)	15515 (58,2)	75,5	57,6	44,7	29,6
<i>Faixa Etária</i> ^{(a) (b)}						
18 a 19	18115 (15,7)	10457 (57,7)	75,0	59,0	45,3	30,8
20 a 29	43990 (38,1)	24049 (54,7)	78,0	61,8	48,9	33,2
30 a 39	28275 (24,5)	16783 (59,4)	77,3	58,6	45,0	30,2
40 a 49	17634 (15,3)	10389 (58,9)	77,3	58,2	44,7	30,9
50 a 59	7088 (6,1)	4114 (58,0)	75,7	57,1	44,9	33,7
60 a 65	408 (0,3)	149 (36,5)	82,0	70,6	61,4	55,6
<i>Estado Civil</i> ^(a)						
Casado	51323 (44,8)	31157 (60,7)	76,6	58,2	44,3	27,7
Divorciado	5524 (4,8)	3458 (62,6)	74,0	54,8	41,6	27,3
Outros	2152 (1,9)	1331 (61,8)	72,6	54,5	41,0	25,0
Solteiro	54187 (47,3)	28801 (53,2)	78,1	61,9	49,2	36,4
Viúvo	1358 (1,2)	810 (59,6)	74,8	55,2	42,9	33,0
<i>Tipo Sanguíneo</i> ^(a)						
A	43299 (37,6)	24475 (56,5)	77,8	60,4	47,1	32,5
AB	4253 (3,7)	2517 (59,2)	74,9	57,6	44,1	31,4
B	12909 (11,2)	7428 (57,5)	76,8	59,0	45,7	31,8
O	54549 (47,4)	31458 (57,7)	76,7	59,3	46,1	31,0
<i>Fator Rh</i> ^(a)						
Negativo	14559 (12,7)	8976 (61,6)	74,1	55,2	41,8	27,3
Positivo	100402 (87,3)	56883 (56,6)	77,5	60,4	47,0	32,3
<i>Unidade</i> ^(b)						
Coleta Externa	58156 (50,3)	33452 (57,7)	80,1	61,5	47,4	32,3
Unidades Pequenas	7311 (6,3)	4435 (60,7)	65,5	49,6	38,6	27,9
Unidades Grandes	50086 (43,3)	28072 (56,0)	75,3	59,3	46,7	32,1

(a) Variáveis que contêm dados *missings*.

(b) Variáveis obtidas na primeira doação.

(c) Dados sumariados em frequência absoluta e percentagem do total de doadores correspondentes em cada grupo.

(d) Dados sumariados em frequência absoluta e percentagem do número de doadores que retornaram em cada grupo.

Figura 4.4.5. Função de sobrevivência representando a propensão à não retornar dos doadores voluntários segundo as variáveis sócio-demográficas. (1- Sexo, 2- Cor da Pele, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh).

Com relação ao fator Rh, doadores com Rh negativo apresentam um maior número de retornos, cerca de 61,6% e também as menores taxas de não retorno em todo o período estudado quando comparado com doadores Rh positivo (56,6%). De maneira geral, independente do tipo sanguíneo do doador, doadores portadores de Rh negativo apresentam uma propensão maior de retornar do que doadores portadores de Rh positivo (curvas não mostradas).

A Figura 4.4.6 apresenta uma aproximação da função risco para cada uma das variáveis sócio-demográficas e o seu comportamento reflete o mesmo impacto que cada variável causou nas respectivas funções de sobrevivências descritas acima.

De maneira geral, a chance instantânea do doador de sangue retornar de ambos os sexos são aproximadamente similares. Isso também é válido para o tipo sanguíneo do doador.

Doadores declarados da cor branca e negra apresentam aproximadamente o mesmo risco à retornar em todo o período estudado, enquanto que doadores declarados da cor amarela apresenta, na maioria das vezes, uma chance de retornar menor.

Exceto a faixa etária dos 60 a 65 anos, a chance instantânea à retornar das demais faixas etárias são aproximadamente as mesmas em todo o período estudado. A faixa etária dos 60 a 65 anos apesar de apresentar o mesmo comportamento decrescente com o passar do tempo, a chance de retornar está sempre abaixo das demais faixas.

A chance à retornar dos doadores de quaisquer estado civil apresenta um comportamento muito similar destacando porém a maior chance à retornar entre os divorciados ou da categoria outros no período de 90 a 120 dias e a menor chance a retornar entre os doadores solteiros no mesmo período.

Doadores portadores de Rh negativo apresentam um risco a retornar sempre maior que os doadores portadores de Rh positivo em todo o período estudado.

Figura 4.4.6. Função Risco instantâneo à retornar dos doadores voluntários segundo as variáveis sócio-demográficas. (1- Sexo, 2- Cor, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh).

Na Figura 4.4.7. estão um resumo do número de doadores que retornaram (em %) em cada uma das categorias de cada variável sócio-demográfica (Tabela 4.4.1.). O número de doadores que retornaram para sua segunda doação, independente de qualquer variável sócio-demográfica, corresponde a uma linha de referência. Doadores declarados como amarelos, pertencentes as faixas etárias dos 20 aos 29 anos ou 60 anos ou mais e solteiros são as categorias que mais se destacam por estarem muito abaixo dessa linha.

Figura 4.4.7. Porcentagem de retorno das categorias de cada variável sócio - demográfica. A notação de cada categoria corresponde as suas iniciais. A linha de referência corresponde ao número de doadores que retornaram independente de qualquer variável sócio-demográfica.

4.5. Aplicação de Modelos de Análise de Sobrevida utilizando o tempo entre a primeira e a segunda doação dos Doadores Voluntários

Para descrever o comportamento de retorno dos doadores voluntários a partir de um modelo probabilístico foram construídos vários modelos paramétricos assumindo diferentes distribuições de probabilidade para o tempo entre a primeira e a segunda doação, de acordo com o comportamento da função risco vista anteriormente.

Foram escolhidas as distribuições de probabilidade Log-Normal e Log-Logística dadas respectivamente em (3.17) e (3.22) pela forma da função risco (apresenta um comportamento crescente até certo ponto e depois decrescente semelhante à função risco obtida empiricamente pelo método atuarial (Figura 4.4.2.)). A distribuição Weibull, dada em (3.14), apesar da sua função risco apresentar um comportamento apenas decrescente, foi escolhida apenas para estudar o desempenho do seu ajuste. A distribuição Gama Generalizada Estendida, dada em (3.28), que inclui várias distribuições de probabilidade como casos particulares tais como a Weibull e a Log-Normal, foi um modelo adicional.

Os ajustes dos modelos paramétricos e paramétricos de longa-duração foram realizados no *software* SAS 9.0 através do procedimento NLP. No PROC NLP, foi utilizado o método de otimização *Trust-Region Optimization* (TRUREG) para encontrar as respectivas estimativas dos parâmetros através da maximização do logaritmo da função de verossimilhança dadas em (3.34) e (3.45). Esse método requer apenas a especificação da função objetivo, no caso, o logaritmo da função de verossimilhança encontrando o valor de máximo local da função.

Para certificar as devidas especificações das funções de verossimilhança de cada uma das distribuições de probabilidade escolhidas, foi ajustado primeiramente os respectivos modelos paramétricos através do procedimento PROC LIFEREG do *software*

SAS 9.0. No PROC LIFEREG, a parametrização das distribuições Weibull, Log-Normal, Log-Logística e Gama Generalizada Estendida (GGE) são dadas na Tabela 4.5.1. Posteriormente, a mesma parametrização utilizada neste procedimento, foi utilizado para a implementação dos ajustes dos modelos paramétricos e paramétricos de longa-duração, utilizando suas estimativas como valores iniciais no PROC NLP para início do algoritmo.

Tabela 4.5.1. Parametrização das distribuições 1)Weibull, 2) Log-Logística, 3)Log-Normal e 4)Gama Generalizada Estendida.

Distribuição	Função Densidade de Probabilidade	Parametrização
1- Weibull	$f(t) = \lambda \gamma (\lambda t)^{\gamma-1} e^{-(\lambda t)^\gamma}$	$\lambda = e^{(-\mu)}$ e $\gamma = \frac{1}{\sigma}$ Onde: $\mu = \alpha_0 + \sum_{j=1}^k \alpha_k x_k$
2- Log-Logística	$f(t) = \frac{\lambda \gamma t^{\gamma-1}}{(1 + \lambda t^\gamma)^2}$	$\lambda = e^{(-\frac{\mu}{\sigma})}$ e $\gamma = \frac{1}{\sigma}$ Onde: $\mu = \alpha_0 + \sum_{j=1}^k \alpha_k x_k$
3- Log-Normal	$f(t) = \frac{1}{\sqrt{2\pi\sigma t}} \exp\left[-\frac{1}{2\sigma^2} \ln(\lambda t)^2\right]$	$\lambda = e^{(-\mu)}$ Onde: $\mu = \alpha_0 + \sum_{j=1}^k \alpha_k x_k$
4- GGE	$f(t) = \frac{ \gamma \alpha^\alpha \lambda^{\alpha\gamma}}{\Gamma(\alpha)} t^{\alpha\gamma-1} \exp[-\alpha(\lambda t)^\gamma]$	$\lambda = e^{(-\mu)}$, $\gamma = \frac{\varphi}{\sigma}$ e $\alpha = \frac{1}{\varphi^2}$ Onde: $\mu = \alpha_0 + \sum_{j=1}^k \alpha_k x_k$

Na Tabela 4.5.2 estão os resultados das estimativas dos parâmetros dos quatro modelos ajustados assumindo as distribuições: 1- Weibull, 2- Log-Logística, 3- Log-Normal e 4- Gama Generalizada Estendida (GGE), considerando previamente uma proporção de doadores improváveis de retornar nula (obtenção de um modelo paramétrico, onde $p = 0$ e função de sobrevivência dada em (3.32)). Na Figura 4.5.1. estão os respectivos ajustes dos modelos propostos. De acordo com o critério BIC, o melhor modelo ajustado é aquele que assume uma distribuição GGE (menor valor) (Tabela 4.5.2.).

Tabela 4.5.2. Estimativas dos parâmetros dos modelos paramétricos ajustados assumindo as distribuições de 1)Weibull, 2) Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida juntamente com os valores obtidos para o Log-Verossimilhança e critério BIC.

	Estimativa	Erro Padrão	Log-Verossimilhança	BIC ^(a)
I - Modelos Paramétricos				
a) Weibull				
μ	7,231	0,0053	-534271,7	1068566,7
σ	1,322	0,0042		
b) Log-Logística				
μ	6,524	0,0053	-527069,4	1054162,1
σ	0,940	0,0030		
c) Log-Normal				
μ	6,601	0,0052	-524213,1	1048449,5
σ	1,525	0,0046		
d) GEE				
μ	4,961	0,0064	-507889,9	1015814,8
σ	0,790	0,0047		
φ	-3,521	0,0268		

(a) Bayesian Information Criterion.

A curva de sobrevivência obtida através do ajuste do modelo assumindo uma distribuição Weibull, não conseguiu estimar a queda acentuada no início do período de seguimento, superestimando a propensão a não retornar, depois, devido o grande número de censuras (doadores que não retornaram para sua segunda doação) a partir de um certo período, o modelo subestimou a propensão a não retornar admitindo um comportamento decrescente acentuado (Figura 4.5.1. – painel superior esquerdo).

As curvas de sobrevivência obtidas através dos ajustes dos modelos assumindo as distribuições Log-Logística e Log-Normal apresentaram um comportamento muito semelhante. Apesar do ganho na estimação no início do período de seguimento diminuindo a sua superestimação, a subestimação no final do seguimento foi similar ao ajuste do modelo assumindo uma distribuição de Weibull (Figura 4.5.1. – painel superior direito e painel inferior esquerdo respectivamente para as distribuições Log-Logística e Log-Normal).

A curva de sobrevivência obtida através do ajuste do modelo assumindo uma distribuição GGE apresenta um comportamento muito próximo do estimado empiricamente pelo método de Kaplan-Meier, principalmente no início do seguimento. A estimação da propensão a não retornar no final do seguimento (após aproximadamente quatro anos) é abaixo do esperado, seguindo a tendência decrescente (Figura 4.5.1. – painel inferior direito).

Figura 4.5.1. Curvas de sobrevivência estimadas pelos modelos paramétricos assumindo as distribuições: 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida seguida da curva estimada por Kaplan-Meier.

Nota-se que a partir daí, que o decréscimo da propensão a não retornar empírica não é mais constante, mas com tendência a estabilizar muito lentamente. Tomado este período como ponto de corte, temos que a taxa de não retorno estimada empiricamente após quatro anos de seguimento é de aproximadamente 37,8%, já no final do seguimento essa taxa cai para 31,8%, o que corresponde num aumento de apenas 6% do número de doadores retornarem em cinco anos.

A taxa de não retorno estimada pelo modelo GGE nos mesmos períodos é de 37,1% e 27,6% respectivamente.

Para avaliar o ajuste dos modelos foram construídos gráficos das estimativas das sobrevivências obtidas por Kaplan-Meier e aquelas advindas de cada modelo (Figura 4.5.2.). A combinação dos pares das estimativas deverá estar aproximadamente ao longo de uma reta para que o modelo ajustado possa ser considerado satisfatório. Nesse caso, apenas o modelo assumindo uma distribuição GGE foi considerado aceitável (Figura 4.5.2 – painel inferior direito).

Para verificar o desempenho do ajuste da classe de modelos de longa-duração, foram construídos vários modelos assumindo as mesmas distribuições utilizadas anteriormente e considerando previamente uma proporção de doadores que não irão retornar não nula (obtenção do modelo paramétrico de longa-duração, onde $0 < p < 1$ e função de sobrevivência dada em (3.39)).

Na Tabela 4.5.3. estão os resultados das estimativas dos parâmetros dos quatro modelos da classe de longa-duração (MLD) ajustados: 1- MLD Weibull, 2- MLD Log-Logística, 3- MLD Log-Normal e 4- MLD Gama Generalizada Estendida (GGE). Na Figura 4.5.3 estão os respectivos ajustes dos modelos propostos. De acordo com o critério BIC, o melhor modelo de longa-duração ajustado é aquele assumindo uma distribuição GGE (menor valor) (Tabela 4.5.3.).

A curva de sobrevivência obtida através do ajuste do modelo paramétrico de longa-duração assumindo uma distribuição Weibull, teve um maior desempenho quando comparado com o respectivo modelo paramétrico (valores de BIC iguais a 1044347,0 e 1068566,7 respectivamente), porém, em nenhum momento a estimação da propensão a não retornar foi próxima daquela estimada empiricamente pelo método de Kaplan-Meier.

Figura 4.5.2. Gráficos das funções de sobrevivência estimadas por Kaplan-Meier *versus* as funções de sobrevivência estimadas pelos modelos assumindo uma distribuição 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida.

Tabela 4.5.3. Estimativas dos parâmetros dos modelos paramétricos de longa-duração ajustados assumindo as distribuições de 1)Weibull, 2)Log-Logística, 3)Log-Normal e 4)Gama Generalizada Estendida juntamente com os valores obtidos para o Log-Verossimilhança e critério BIC.

	Estimativa	Erro Padrão	Log-Verossimilhança	BIC ^(a)
I - Modelos Paramétricos de Longa-Duração				
a) Weibull			-522156,0	1044347,0
μ	6,114	0,0040		
σ	0,859	0,0026		
β_0	0,648	0,0073		
b) Log-Logística			-514976,5	1029988,0
μ	5,629	0,0039		
σ	0,515	0,0019		
β_0	0,639	0,0075		
c) Log-Normal			-513499,4	1027033,8
μ	5,680	0,0038		
σ	0,864	0,0029		
β_0	0,639	0,0074		
d) GEE			-507452,9	1014952,4
μ	5,006	0,0070		
σ	0,709	0,0408		
φ	-2,528	0,0044		
β_0	1,408	0,0327		

(a) Bayesian Information Criterion.

A proporção estimada de doadores improváveis de retornar foi de 34,3% (Figura 4.5.3. – painel superior esquerdo).

Assim como o caso anterior, as curvas de sobrevivência obtidas através dos ajustes dos modelos paramétricos de longa-duração assumindo tanto a distribuição Log-Logística quanto a distribuição Log-Normal apresentou um comportamento muito semelhante e também mais satisfatório que os respectivos modelos paramétricos. Para os modelos assumindo uma distribuição Log-Logística, os valores do BIC foram 1029988,0 e 1054162,1 respectivamente. Já para o modelo assumindo uma distribuição Log-Normal, os valores do BIC foram 1027033,8 e 1048449,5 respectivamente (Figura 4.5.3. – painel superior direito e painel inferior esquerdo respectivamente).

De uma maneira geral, ambos os modelos apresentaram ser satisfatoriamente bem ajustados principalmente no início do intervalo. As proporções estimadas de doadores que não irão retornar para uma segunda doação para as duas distribuições foram próximas, cerca de 34,5%.

Figura 4.5.3. Curvas de sobrevivência estimadas pelos modelos de longa-duração assumindo as distribuições de 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida, seguida da curva estimada por Kaplan-Meier.

A curva de sobrevivência obtida através do ajuste do modelo paramétrico de longa-duração assumindo uma distribuição GGE apresenta um comportamento muito semelhante ao modelo paramétrico ajustado anteriormente (BIC iguais a 1014952,4 e 1015814,8 respectivamente), porém a utilização de um parâmetro adicional fez com que a curva caísse mais suavemente e ajustasse melhor no final do seguimento. Suas estimativas são muito próximas do estimado empiricamente pelo método de Kaplan-Meier.

A proporção estimada de doadores que não irão retornar para uma segunda doação foi de 19,6%, o que não reflete no período estudado. Isso indica que, seguindo essa tendência decrescente, é esperado ainda um retorno de aproximadamente 12,2% após o período aqui estudado.

Para avaliar o ajuste dos modelos ajustados foram construídos gráficos das estimativas das sobrevivências obtidas por Kaplan-Meier e aquelas advindas de cada modelo de longa-duração (Figura 4.5.4.). Nesse caso, apesar dos modelos assumindo uma distribuição Log-Logística e Log-Normal estarem relativamente próximas do esperado, apenas o modelo assumindo uma distribuição GGE foi considerado aceitável.

A partir daqui, admitindo, pois o modelo paramétrico de longa-duração com distribuição GGE como o melhor modelo ajustado, a análise segue incluindo neste modelo as variáveis relacionadas com o perfil do doador (covariáveis) e assim verificar o seu efeito no comportamento de retorno através da função de sobrevivência.

O ajuste dos modelos paramétricos de longa-duração considerando as covariáveis foi realizado através da biblioteca *gfcure* através do *software* R. O programa *gfcure* apresenta uma ampla flexibilidade em trabalhar com diversas distribuições. Ele divide a saída em duas partes: uma referente ao componente paramétrico que modela o tempo entre a primeira e a segunda doação e outra referente ao componente logístico que modela a proporção de doadores improváveis de retornar, p .

Figura 4.5.4. Gráficos das funções de sobrevivência estimadas por Kaplan-Meier *versus* as funções de sobrevivência estimadas pelos modelos de longa-duração assumindo uma distribuição 1)Weibull, 2)Log-Logística, 3)Log-Normal, e 4)Gama Generalizada Estendida.

4.6. Aplicação de Modelos de Análise de Sobrevivência utilizando o tempo entre a primeira e a segunda doação dos Doadores Voluntários e covariáveis

Para investigar o efeito das covariáveis relacionadas com o perfil do doador (listadas na Tabela 4.3.1.) e o tempo entre a primeira e a segunda doação, ajustamos modelos de regressão de longa-duração, baseados nas distribuições: 1) Weibull, 2) Log-Logística, 3) Log-Normal e 4) GGE, descritas na seção anterior, com função de sobrevivência dada em (3.41). Nesse caso, partimos do pressuposto que os doadores prováveis e improváveis de retornar para uma segunda doação apresentam perfis diferentes. Isso significa que, no modelo proposto, o efeito de cada covariável relacionada com o seu perfil irá afetar tanto o parâmetro de locação (componente paramétrico), λ , quanto à proporção p (componente logístico). Os resultados para esses ajustes foram obtidos através da biblioteca *gfcure* (PENG; DEAR; DENHAM, 1998) implementada no *software* R.

Nas Figuras 4.6.1 e 4.6.2 estão os gráficos utilizados para checar a suposição de proporcionalidade de riscos, ou seja, do $\ln[-\ln S_{KM}(t)]$ versus t , o que equivale a partir da equação (3.11), ao gráfico $\ln[H(t)]$ versus t . Note que em todos os casos, essa suposição não é satisfeita, sendo um caso onde o ajuste de um modelo de mistura pode ser apropriado.

Em seguida, ajustamos modelos de regressão de longa-duração considerando cada uma das distribuições citadas e uma covariável de cada vez. Consideramos o parâmetro de locação, λ , e o parâmetro, p , relacionados à covariável em questão. Cada covariável foi transformada em uma variável indicadora, assumindo valores 0 (categoria de referência) ou 1. Em cada ajuste, construímos gráficos das funções estimadas por Kaplan-Meier versus as funções estimadas pelos modelos (resultados não mostrados).

Figura 4.6.1. Logaritmo da função risco instantâneo à retornar acumulado dos doadores voluntários segundo unidade de coleta.

Na Tabela 4.6.1 estão os resultados obtidos assumindo uma distribuição GGE. Na Tabela 4.6.2 está o cálculo das respectivas proporções esperados de doadores improváveis de retornar. Tais resultados podem ser visualizados na Figura 4.6.3.

Em todos os casos, o modelo assumindo uma distribuição GGE foi o mais satisfatório. E ainda, comparando os ajustes dos modelos entre as quatro diferentes distribuições de probabilidade, notamos, que o critério BIC é menor quando utilizada a distribuição GGE (Tabela 4.6.3).

Em um segundo passo, ajustamos modelos de regressão de longa-duração considerando todas as covariáveis simultaneamente. Novamente consideramos o parâmetro de locação, λ , e o parâmetro responsável pela fração de doadores improváveis de retornar, p , relacionados a um vetor de covariáveis, x e z presentes na equação (3.41).

Figura 4.6.2. Logaritmo da função risco instantâneo à retornar acumulado dos doadores voluntários segundo as variáveis sócio-demográficas. (1- Sexo, 2- Cor, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh).

Tabela 4.6.1. Estimativas dos parâmetros do modelo de regressão de longa-duração com distribuição para o tempo entre a primeira e a segunda doação GGE para cada covariável relacionada com o perfil do doador voluntário – Ajuste I.

		Ajuste (I) ^(a)		
		Intercepto	Estimativa	Erro Padrão
Parâmetros obtidos no componente paramétrico				
1) Sexo	(Masculino vs Feminino)	5,023	-0,158	0,005
2) Cor da Pele	(Branca vs Amarela)	5,045	-0,039	0,034
	(Negra vs Amarela)		-0,086	0,034
3) Faixa etária	(18 a 19 vs 60 a 65)	4,986	-0,052	0,046
	(20 a 29 vs 60 a 65)		-0,003	0,046
	(30 a 39 vs 60 a 65)		0,030	0,046
	(40 a 49 vs 60 a 65)		0,048	0,046
	(50 a 59 vs 60 a 65)		0,025	0,047
4) Estado Civil	(Casado vs Solteiro)	4,977	0,035	0,005
	(Divorciado vs Solteiro)		0,013	0,011
	(Outros vs Solteiro)		-0,060	0,017
	(Viúvo vs Solteiro)		0,054	0,022
5) Tipo Sangüíneo	(AB vs A)	5,003	-0,054	0,013
	(B vs A)		-0,018	0,008
	(O vs A)		-0,009	0,005
6) Fator Rh	(Negativo vs Positivo)	4,998	-0,020	0,007
7) Unidade de Coleta	(Unidades Pequenas vs Coleta Externa)	5,079	-0,273	0,009
	(Unidades Grandes vs Coleta Externa)		-0,163	0,005
Parâmetros obtidos no componente logístico				
1) Sexo	(Masculino vs Feminino)	1,678	-0,035	0,032
2) Cor da Pele	(Branca vs Amarela)	0,614	0,835	0,119
	(Negra vs Amarela)		0,953	0,121
3) Faixa etária	(18 a 19 vs 60 a 65)	0,238	1,245	0,15
	(20 a 29 vs 60 a 65)		1,039	0,148
	(30 a 39 vs 60 a 65)		1,371	0,149
	(40 a 49 vs 60 a 65)		1,361	0,151
	(50 a 59 vs 60 a 65)		1,217	0,155
4) Estado Civil	(Casado vs Solteiro)	1,164	0,578	0,029
	(Divorciado vs Solteiro)		0,777	0,081
	(Outros vs Solteiro)		0,851	0,139
	(Viúvo vs Solteiro)		0,430	0,129
5) Tipo Sangüíneo	(AB vs A)	1,400	0,151	0,073
	(B vs A)		0,074	0,044
	(O vs A)		0,109	0,028
6) Fator Rh	(Negativo vs Positivo)	1,422	0,468	0,05
7) Unidade de Coleta	(Unidades Pequenas vs Coleta Externa)	1,410	0,346	0,064
	(Unidades Grandes vs Coleta Externa)		-0,119	0,025

(a) Ajuste de cada covariável individualmente.

Figura 4.6.3. Função de sobrevivência representando a propensão à não retornar dos doadores de sangue voluntários de primeira vez segundo as variáveis sócio-demográficas e os respectivos ajustes pelo modelo de regressão de longa duração com distribuição GGE. (1- Sexo, 2- Cor, 3- Faixa Etária, 4- Estado Civil, 5- Tipo Sanguíneo, 6- Fator Rh).

Figura 4.6.4. Função de sobrevivência representando a propensão à não retornar dos doadores de sangue voluntários de primeira vez segundo unidade de coleta e respectivo ajuste pelo modelo de longa-duração com distribuição GGE.

Desta maneira, ajustamos um modelo de regressão para cada uma das quatro distribuições aqui apresentadas. Novamente o modelo baseado na distribuição GGE apresentou evidências de melhor ajuste, apresentando um valor para o critério BIC menor que o observado para as demais distribuições (valores do critério BIC igual a 276467,7, contra 292548,8 para o modelo assumindo uma distribuição Log-Logística e 289738,8 para o modelo assumindo uma distribuição Log-Normal) (Tabela 4.6.3).

Na Tabela 4.6.4 estão os resultados obtidos assumindo uma distribuição GGE. Estimativas dos parâmetros com sinal negativo no componente paramétrico indicam uma propensão a não retornar menor do que a categoria de referência. Estimativas dos parâmetros negativas no componente logístico indicam uma proporção esperada maior de doadores improváveis de retornar¹⁷.

¹⁷ Note aqui que, para interpretar a magnitude do parâmetro associado ao componente paramétrico paralelo as simulações feitas na seção Sujeito e Métodos é necessário considerar que o parâmetro μ está em função de λ .

Tabela 4.6.2. Proporção esperada doadores voluntários improváveis de retornar de acordo com as covariáveis relacionadas com o seu perfil.

Variáveis		<i>p</i>
1) Sexo	Feminino	0,157
	Masculino	0,162
2) Cor da Pele	Amarela	0,351
	Branca	0,190
	Negra	0,173
3) Faixa etária	18 a 19	0,185
	20 a 29	0,218
	30 a 39	0,167
	40 a 49	0,168
	50 a 59	0,189
	60 a 65	0,441
4) Estado Civil	Casado	0,149
	Divorciado	0,126
	Outros	0,118
	Solteiro	0,238
	Viúvo	0,169
5) Tipo Sangüíneo	A	0,198
	AB	0,175
	B	0,186
	O	0,181
6) Fator Rh	Negativo	0,131
	Positivo	0,194
7) Unidade de Coleta	Unidades Pequenas	0,147
	Unidades Grandes	0,216
	Coleta Externa	0,196

De maneira geral, observando as estimativas dos parâmetros relacionados com o componente paramétrico, verificamos que apenas as variáveis: sexo e unidade de coleta são as covariáveis que apresentam os maiores efeitos, dada a magnitude de sua estimativa.

Assim, entre os doadores prováveis de retornar, doadores do sexo masculino tendem a retornar mais cedo do que doadores do sexo feminino.

Tabela 4.6.3. Valores do critério BIC para os modelos ajustados assumindo uma distribuição 1)Weibull, 2)Log-Logística, 3)Log-Normal e 4)GGE para o tempo entre a primeira e a segunda doação.

Ajustes	Weibull ^(b)	Log-Logística ^(b)	LogNormal ^(b)	GGE ^(b)
<i>Modelo Paramétrico</i>				
Sem covariáveis	337014,1	322609,5	316930,3	284266,0
Com covariáveis	331310,4	316915,4	311438,2	277934,3
<i>Modelo Paramétrico Longa-Duração</i>				
Sem covariáveis	312794,6	298435,4	295514,6	283404,8
Apenas sexo	312757,7	298420,5	295497,1	282507,9
Apenas cor da pele	^(a)	296994,0	294082,6	281971,8
Apenas faixa etária	^(a)	^(a)	295303,9	283079,4
Apenas estado civil	309847,7	295584,1	292691,1	280661,4
Apenas tipo sanguíneo	311801,3	297456,9	294544,3	282466,5
Apenas fator Rh	311500,9	297156,1	294249,1	282209,9
Apenas unidade de coleta	312151,6	297434,6	294576,6	281698,7
Com covariáveis e p fixo	307779,9	293317,3	290498,7	277189,9
Com covariáveis	^(a)	292548,8	289738,8	276467,7

(a) Variáveis no qual houve problemas na estimação dos parâmetros e/ou erros-padrão.

(b) Valores do critério BIC obtidos a partir da maximização do logaritmo da função de verossimilhança considerando o logaritmo do tempo. A biblioteca *gfcure* maximiza o log-verossimilhança usando o logaritmo do tempo de sobrevivência como variável resposta. Este log-verossimilhança, difere daquele calculado com o tempo de sobrevivência propriamente dito, porém, não afeta as estimativas dos parâmetros e dos erros-padrão.

Dado que o intervalo mínimo entre as doações difere entre ambos os sexos, esse fenômeno pode ser justificado. Note pelo gráfico do logaritmo da função risco (Figura 4.6.2 – painel superior esquerdo) que essa diferença na propensão a retornar corresponde a esse período.

Doadores provenientes da coleta interna também tendem a ser mais propensos a retornar do que aqueles provenientes da coleta externa.

As demais covariáveis estão associadas parâmetros com estimativas de magnitudes relativamente pequenas, próximas à zero, evidenciando que não estão relacionados com o tempo de retorno do doador.

Tabela 4.6.4. Estimativas dos parâmetros do modelo de regressão de longa-duração com distribuição para o tempo entre a primeira e a segunda doação GGE para cada covariável relacionada com o perfil do doador voluntário – Ajuste II.

		Ajuste (II) ^(a)	
		Estimativa	Erro Padrão
Parâmetros obtidos no componente paramétrico			
Intercepto		5,090	
1) Sexo	(Masculino vs Feminino)	-0,131	0,005
2) Cor da Pele	(Branca vs Amarela)	0,026	0,031
	(Negra vs Amarela)	-0,007	0,031
3) Faixa etária	(18 a 19 vs 60 a 65)	-0,067	0,041
	(20 a 29 vs 60 a 65)	-0,024	0,041
	(30 a 39 vs 60 a 65)	-0,007	0,041
	(40 a 49 vs 60 a 65)	0,002	0,041
	(50 a 59 vs 60 a 65)	-0,021	0,041
4) Estado Civil	(Casado vs Solteiro)	0,003	0,005
	(Divorciado vs Solteiro)	-0,045	0,010
	(Outros vs Solteiro)	0,001	0,016
	(Viúvo vs Solteiro)	-0,06	0,020
5) Tipo Sangüíneo	(AB vs A)	-0,034	0,012
	(B vs A)	-0,011	0,007
	(O vs A)	-0,006	0,005
6) Fator Rh	(Negativo vs Positivo)	-0,019	0,006
7) Unidade de Coleta	(Unidades Pequenas vs Coleta Externa)	-0,226	0,009
	(Unidades Grandes vs Coleta Externa)	-0,136	0,005
Parâmetros obtidos no componente logístico			
Intercepto		-1,170	
1) Sexo	(Masculino vs Feminino)	-0,061	0,031
2) Cor da Pele	(Branca vs Amarela)	0,761	0,124
	(Negra vs Amarela)	0,929	0,128
3) Faixa etária	(18 a 19 vs 60 a 65)	1,933	0,171
	(20 a 29 vs 60 a 65)	1,513	0,167
	(30 a 39 vs 60 a 65)	1,534	0,168
	(40 a 49 vs 60 a 65)	1,387	0,169
	(50 a 59 vs 60 a 65)	1,222	0,173
4) Estado Civil	(Casado vs Solteiro)	0,775	0,039
	(Divorciado vs Solteiro)	0,972	0,093
	(Outros vs Solteiro)	0,992	0,158
	(Viúvo vs Solteiro)	0,708	0,146
5) Tipo Sangüíneo	(AB vs A)	0,245	0,085
	(B vs A)	0,086	0,049
	(O vs A)	0,113	0,031
6) Fator Rh	(Negativo vs Positivo)	0,465	0,054
7) Unidade de Coleta	(Unidades Pequenas vs Coleta Externa)	0,473	0,086
	(Unidades Grandes vs Coleta Externa)	-0,149	0,030

(a) Ajuste de todas as covariáveis simultaneamente.

Quando observamos as estimativas dos parâmetros relacionados com o componente logístico, verificamos que as covariáveis cor da pele, faixa etária, estado civil, tipo sanguíneo, fator Rh e unidade de coleta apresentam os maiores efeitos, dada a magnitude da sua estimativa.

Assim, é esperado que doadores declarados como negros ou brancos tendem a ter uma proporção de não retorno menor do que doadores declarados da cor amarela. Com relação à faixa etária, é esperado uma proporção de doadores improváveis de retornar a medida que esta aumenta. O *status* estado civil evidencia uma considerável proporção de improváveis de retornar entre os doadores solteiros. Doadores portadores de fator Rh negativo apresentam uma proporção de não retorno esperada menor quando comparado com doadores com fator Rh positivo, bem como aqueles portadores do tipo sanguíneo AB e O em relação aos tipos A e B.

É esperado ainda um número menor de doadores que não irão retornar entre as unidades pequenas.

O ajuste do modelo (II) indica que cerca de 2% dos doadores voluntários do sexo feminino, cor da pele negra, pertencente a faixa etária de 18 a 19 anos, tendo como status de estado civil outros, com tipo sanguíneo AB, portadora de fator Rh negativo e proveniente de uma das Unidades Pequenas, não irão retornar para uma nova doação. Além disso, estima que aproximadamente 80% dos indivíduos do sexo masculino, cor da pele amarela, pertencente à faixa etária dos 60 a 65 anos, solteiro, com tipo sanguíneo A, portador de fator Rh positivo e proveniente de uma das Unidades Grandes, também não irão retornar.

Em termos comparativos, a inclusão das covariáveis no modelo de regressão de longa-duração possibilita explicar o tempo até o retorno do doador (valor do critério BIC menor quando comparado com o modelo sem covariáveis – valores iguais a 283404,8 e 276467,7 respectivamente para os modelos sem covariáveis e com a inclusão de covariáveis).

Esse modelo também se mantém mais satisfatório quando comparado com o modelo de longa-duração com p fixo, ou seja, sem a inclusão das covariáveis na proporção, p , (valor do critério BIC igual a 277189,9) (Tabela 4.6.5.).

Propomos ainda um outro modelo apenas com as covariáveis que tiveram um impacto substancial no componente paramétrico (sexo e unidade de coleta) e/ou no componente logístico (cor da pele, faixa etária, estado civil, tipo sanguíneo e fator RH) (Ajuste III – modelo reduzido) obtendo um valor de BIC muito próximo do modelo completo com uma substancial diminuição no número de parâmetro (valores do critério BIC iguais a 276472,0 e 276467,7 respectivamente para os modelos reduzido e completo).

Ainda que o critério BIC indique os ajustes (II) e (III) como os modelos mais satisfatórios na explicação do comportamento de retorno do doador através das variáveis relacionadas com o seu perfil, é necessário verificar o quão eles são capazes de indicar o perfil do doador improvável de retornar, ou seja, validar o modelo.

Tabela 4.6.5. Valores da maximização do Log-Verossimilhança e do critério BIC dos modelos assumindo uma distribuição GGE para o tempo entre a primeira e a segunda doação.

	Log-Verossimilhança ^(a)	Nº de Parâmetros	BIC
<i>Modelo Paramétrico</i>			
Sem covariáveis	-142115,5	3	284266,0
Com covariáveis	-138844,9	21	277934,3
<i>Modelo Paramétrico Longa-Duração</i>			
Sem covariáveis	-141679,1	4	283404,8
Com covariáveis e p fixo	-138466,9	22	277189,9
Com covariáveis (II)	-138001,0	40	276467,7
Com covariáveis (III)	-138096,3	24	276472,0

(a) A biblioteca *gfcure* maximiza o log-verossimilhança usando o logaritmo do tempo de sobrevida como variável resposta. Este log-verossimilhança, difere daquele calculado com o tempo de sobrevida propriamente dito, porém, não afeta as estimativas dos parâmetros e dos erros-padrão.

A validação de modelos tem como finalidade verificar o poder discriminatório de um modelo ajustado observando assim os seus valores preditos.

No caso, partimos do pressuposto que os valores estimados da proporção p obtida por um determinado modelo são correlacionados com a não ocorrência do retorno para uma nova doação, de tal maneira que altos valores de p seriam encontrados entre os doadores que não retornaram, indicando o perfil improvável de retornar, da mesma maneira que valores pequenos de p seriam encontrados entre os doadores que retornaram, indicando o perfil do doador provável de retornar. Com isso, a proporção p seria um critério para fazer predições a respeito da ocorrência ou não desse evento.

Essa previsão pode ser feita estabelecendo um ponto de corte em p de forma que valores maiores que este, seriam classificados como possíveis doadores improváveis de retornar e valores menores que este como indivíduos prováveis de retornar.

Todas as possíveis classificações dos indivíduos, variando esses pontos de corte, poderiam ser reunidas a partir da construção da curva ROC (*Receiver Operating Characteristic*) (ZWEIG; CAMPBELL, 1993). A curva ROC é um método gráfico muito utilizado na área médica para medir o desempenho de testes diagnósticos e consiste num gráfico simples dos valores de sensibilidade *versus* um menos a especificidade e permite estudar a variação de ambos para diferentes valores de ponto de corte.

A sensibilidade nesse caso, seria dada pela proporção de doadores classificados como improváveis de retornar dado que eles não retornaram e a especificidade, pela proporção de doadores classificados como prováveis de retornar dado que eles retornaram.

Um modelo que estime p com poder discriminatório, irá apresentar uma curva ROC onde os seus respectivos valores se concentram no canto superior esquerdo do gráfico. A área abaixo da curva ROC está associada ao poder discriminante do modelo, quanto maior for à área, maior é a capacidade do modelo em discriminar os doadores que jamais retornarão.

Na validação, geralmente, a amostra é dividida em duas partes: uma para o desenvolvimento do modelo (amostra de desenvolvimento) e outra para a sua validação (amostra de validação). Usualmente, é destinado cerca de 70% dos dados para a amostra de desenvolvimento (THOMAS; EDELMAN; CROOK, 2002).

As amostras de validação são utilizadas para verificar se o modelo ajustado mantém seu poder discriminatório para outras amostras da mesma população.

A partir de uma amostragem sistemática, 75% dos doadores foram selecionados para a amostra de desenvolvimento. Na Figura 4.6.5. estão os resultados obtidos dos modelos ajustados considerando todas as possíveis combinações das covariáveis relacionadas com o perfil do doador incluídas em ambos os componentes paramétrico e logístico.

Os modelos que apresentaram maiores área sob a curva, cerca de 0,551, continha as covariáveis: 1) estado civil, tipo sanguíneo, fator Rh e unidade de coleta e 2) cor da pele, estado civil, tipo sanguíneo, fator Rh e unidade de coleta, destacando o primeiro modelo citado pelo menor número de covariáveis.

Mesmo assim, ambos os modelos indicam uma baixa capacidade preditiva das covariáveis estudadas em identificar o perfil do doador improvável de retornar para outras amostras.

Figura 4.6.5. Resultados da área sob a curva ROC obtidos com os ajustes dos modelos de acordo com o número de covariáveis inseridas no modelo.

CAPÍTULO V

Discussão

O grande objetivo de um serviço de hemoterapia é garantir estabilidade, segurança e adequada oferta de sangue. Esforços na captação de doadores provenientes de populações de baixo risco vão de encontro a esse objetivo, além da sua fidelização à doação de sangue.

O doador voluntário, candidato à doação movido por sentimentos altruístas, apresenta maior margem de segurança e por isso o estudo do seu comportamento de retorno é fundamental para o direcionamento de estratégias para mantê-lo (LUDWIG; RODRIGUES, 2005).

Assim, utilizamos uma abordagem paramétrica por meio de modelos de longa-duração (MALLER; ZHOU, 1996; PENG; DEAR; DENHAM, 1998; MAZUCHELI; LOUZADA-NETO; ACHCAR, 2003) para estudar o comportamento de retorno do doador de sangue, partindo do pressuposto que entre os candidatos aptos na primeira doação há dois subgrupos de indivíduos: um mais propenso a retornar e outro improvável de retornar. Com isso, podemos obter informações tanto da distribuição dos tempos entre aqueles doadores prováveis de retornar, indicando pontualmente o número de indivíduos que retornarão com o passar do tempo e também estimar a proporção de doadores que jamais retornarão.

A discriminação desses indivíduos é importante num serviço de hemoterapia para o planejamento de estratégias em reter e incentivar doadores a futuras doações.

Porém, tal abordagem não vem sendo utilizada na prática, sendo comumente vista para estudar o comportamento de retorno do doador à utilização de modelos de riscos proporcionais de Cox. Sua popularidade se deve principalmente à facilidade de interpretação e pela sua implementação nos *softwares* usuais. Entretanto, requer que a suposição de proporcionalidade de riscos seja satisfeita, além de se limitar em observar o comportamento de retorno como um todo (SPOSTO, 2002; FRANKEL; LONGMATE, 2002).

Aqui, evidências de que o fator Rh auxilia na explicação do tempo até o retorno do doador foram encontradas, bem como nos estudos de James e Matthews (1993, 1996) entre os doadores com tipo sanguíneo O e Ownby, Watanabe e Nass (1999). Porém, nesses estudos, os autores não distinguem esse efeito de acordo com doadores prováveis e improváveis de retornar.

James e Matthews (1996), por exemplo, ao observar a não proporcionalidade de riscos nas covariáveis, estratificou o período estudado e concluiu que doadores portadores de Rh negativo apresentam sim uma propensão maior de retornar apenas durante um ano decorrido da primeira doação. Entre os primodoadores da região de abrangência de Ribeirão Preto encontramos essa evidência apenas no fim do seguimento o que corresponde à proporção de doadores improváveis de retornar.

Sposto (2002) alerta quanto à utilização de modelos de mistura quando o tempo de seguimento é insuficiente para inferir a presença de indivíduos suscetíveis e não suscetíveis ao evento. Nesse caso, a estimação da fração de imunes é sensível à escolha da distribuição para os tempos estudados.

Nesta casuística, a indicação da distribuição a ser assumida partiu não só pelos critérios adotados, mas também observar a subestimação da propensão à não retornar no final do seguimento sugerido pelo modelo paramétrico usual. Além disso, nesse contexto, é razoável analisar os dados considerando uma classe de modelos que incorpore a existência de duas populações possivelmente heterogêneas.

Com esse mesmo intuito, poderíamos utilizar uma abordagem semi-paramétrica para a construção de modelos de longa-duração (PENG; DEAR, 2000), ou seja, uma generalização do modelo de Cox tradicionalmente utilizada no contexto de doação de sangue, com enfoque em modelos longa-duração.

Porém, quando a distribuição dos tempos é bem especificada, não há grandes diferenças entre essas duas abordagens, exceto quando a função risco diminui rapidamente. Neste caso, as estimativas do modelo semi-paramétrico tende apresentar variações grandes e vieses (PENG; CARRIERE, 2002). Todavia, essa abordagem pode ser útil em situações onde há dificuldade em determinar uma distribuição apropriada para os tempos de falha entre os indivíduos suscetíveis.

A abordagem paramétrica, porém, traz uma maior flexibilidade no estudo da distribuição dos tempos estudados, como visto em Yamaguchi (1992) e Peng, Dear e Denham (1998), que descreveram a utilização da distribuição gama generalizada estendida e da distribuição F. Estas distribuições acomodam outras distribuições entre aqueles indivíduos suscetíveis ao evento, como foi visto no presente estudo ao assumirmos uma distribuição gama generalizada estendida para os tempos até o retorno do doador. Além disso, podemos assumir o efeito de covariáveis nos diferentes parâmetros da respectiva distribuição (MAZUCHELI; LOUZADA-NETO; ACHCAR, 2003).

Além disso, a implementação do algoritmo para estimar seus parâmetros é mais complexa do que os modelos paramétricos. Uma biblioteca implementada em S-Plus, denominada *NPCURE*, para ajustar modelos semiparamétricos já foi desenvolvida, mas requer autorização do autor (PENG; DEAR, 2000) para seu uso. Recentemente, foi desenvolvida uma macro no *software* SAS para o ajuste de modelos de mistura paramétricos e semi-paramétricos, denominada *PSPMCM* (CORBIERE; JOLY, 2007) com uma ampla variedade de distribuições, funções de ligação para estimar a fração de não suscetíveis ao evento, métodos para estimar os erros padrão das estimativas dos parâmetros do modelo de Cox, entre outros.

Aqui, o ajuste de um modelo com distribuição GGE possibilitou estimar que 20% dos candidatos voluntários e aptos à doação não irão retornar. No entanto a interpretação desse parâmetro como a fração esperada de doadores que não retornam a uma nova doação é subjetiva, considerando os dados aqui utilizados. Dado que a curva de sobrevivência tende a p quando $t \rightarrow \infty$, observamos que a curva aproxima-se de 20% somente para valores de t maiores de 10 mil dias (ou 27 anos), um tempo demasiadamente grande quando pretende-se estimar o intervalo até uma nova doação. Observando o ajuste do modelo com fração nula de doadores improváveis de retornar, podemos observar que após aproximadamente quatro anos do início do seguimento, o decréscimo da propensão de não retorno já não é mais constante, tendendo a estabilizar muito lentamente. Tomando este período como ponto de corte, temos que a propensão de não retorno estimada após quatro anos de seguimento é de aproximadamente 37,8%. Já no final do seguimento (nove anos) é de aproximadamente 31,8% o que corresponde a um aumento de retorno dos doadores de apenas 6%.

Este modelo de longa-duração também fornece informações pontuais da frequência esperada de primodoadores que retornam a uma nova doação ao longo do tempo. Observamos no seu componente logístico, valores relativamente altos das estimativas dos parâmetros, levando a uma fração esperada de não retorno diferente dependendo do perfil do doador, e baixos valores no componente paramétrico, levando a uma distribuição de retorno, entre aqueles prováveis de retornar, muito próximas.

Porém, a importância destes valores, não pode ser julgada por testes de hipóteses, dado que a casuística é grande, e, conseqüentemente, pequenos erros padrão levariam à rejeição da hipótese da nulidade dos parâmetros associados às variáveis ainda que não houvesse uma significância prática. Ainda que baixos valores das estimativas impliquem, do ponto de vista estatístico, em diferenças nulas entre dois determinados subgrupos, na prática, o contingente de bolsas de sangue coletadas pode ser significativo, perante o grande

número de doadores de sangue atendidos na região de abrangência do Centro Regional de Hemoterapia de Ribeirão Preto.

Nesse caso, o modelo poderá sim ser capaz de indicar um público pequeno, porém mais propenso a retornar do que os demais dentro de um determinado período.

Nitidamente observamos que doadores do sexo masculino tendem a retornar mais cedo do que doadores do sexo feminino. Esse efeito sugere ser devido ao distinto intervalo mínimo entre as doações para ambos os sexos. Também notamos que doadores provenientes das unidades pequenas tendem a ser mais propensos a retornar do que aqueles provenientes das unidades grandes e coleta externa.

Além disso, é esperada uma proporção de doadores improváveis de retornar maior entre os indivíduos declarados da cor amarela, com sessenta a sessenta e cinco anos, solteiros, bem como entre os portadores de fator Rh positivo e provenientes de unidades grandes ou coleta externa.

Entre os indivíduos declarados da cor amarela, porém, observamos posteriormente que estes tendem a ser solteiros, com menos de 30 anos e um nível de escolaridade entre 2º grau completo/3º grau incompleto ou pós-graduação, o que possivelmente justifica o seu não retorno por ser um público característico de uma região que agrega cidades universitárias e, portanto, estar associado a uma passagem temporária por elas.

Apesar do modelo ser considerado satisfatório em determinar o comportamento de retorno doador, ele se mostrou incapaz de discriminar perfis de doadores improváveis de retornar baseado na proporção p para outras amostras.

Outras considerações, porém, merecem destaque nesta análise, tais como, o controle de fatores externos, como o número de óbitos nesse período e o local onde o doador mora, dos quais não foram considerados. Porém, numa tentativa de verificar este efeito, censuramos, numa segunda etapa, os indivíduos que realizaram sua primeira doação após os

cinquenta e sete anos de idade de acordo com a data limite para a doação, ou seja, sessenta e cinco anos, caso este não havia retornado antes deste período, totalizando cento e vinte e nove indivíduos. Além disso, excluímos os indivíduos que moravam em algum estado que não fizesse divisa com o estado de São Paulo, num total de duzentos e quarenta e sete pessoas. Os resultados obtidos foram bastante semelhantes aos apresentados até então (resultados não mostrados).

O estudo também não propiciou a inclusão de outras variáveis tais como: nível de escolaridade, devido a grande quantidade de não resposta e aprovação na triagem, dado que este estudo foi conduzido apenas com os primodoadores voluntários e indivíduos reprovados temporariamente não respondiam à questão tipo de doação. Evidências da importância dessa variável se encontram no ANEXO G.

Outro fator levantado posteriormente foi à distinção do comportamento de retorno do doador de acordo com o ano da primeira doação, de tal maneira que doadores que iniciaram seu ciclo de doação no início do seguimento tendem a retornar mais cedo do que doadores que iniciaram no fim do seguimento (ANEXO H). Tal observação sugere uma maior investigação, mas que acrescenta a uma série de fatores que envolvem a decisão de retornar, começando pela sua motivação.

Glynn et al. (2002) aponta que independentemente do perfil do doador, a motivação primária em doar é o altruísmo e que o papel das agências de hemoterapia é justamente desenvolver um processo contínuo voltado a esta identidade humanista para que isto se torne um hábito. A indução da família e dos amigos também apresenta um papel importante neste processo, principalmente entre aqueles que doam pela primeira vez.

O hábito em retornar, segundo Suárez et al. (2004), pode ser ainda facilitado pelas condições que o meio proporciona tais como, conveniência, acesso aos postos de coleta, entre outros.

No Brasil, segundo a pesquisa patrocinada pela ANVISA (BRASIL, 2004c), a motivação altruísta é citada por 70% dos doadores e o principal motivo por não doar freqüentemente é justamente a falta de tempo (24%) ou pela falta de oportunidade /distância (10%).

Misje et al. (2005) ressalta a possibilidade de haver uma combinação de motivações que levam um individuo a doar, não só altruístas, mas também razões sociais, de auto-estima, entre outros. Conclui ainda que, em geral, a experiência na doação não revela uma maior motivação do que outra.

Tais resultados reforçam as evidências de que independente do perfil do doador, a motivação intrínseca, seja ela social ou altruísta, é a chave para a sustentação dos bancos de sangue.

CAPÍTULO VI

Conclusão

Os resultados apresentados neste trabalho descrevem o comportamento de retorno dos doadores de sangue que se candidataram ao menos uma vez à doação de sangue em algum serviço de hemoterapia pertencente à área de abrangência de Ribeirão Preto. Entre eles, observamos,

- ⇒ Aumento no número de doações voluntárias e de doadores voluntários ao longo do histórico de doação.
- ⇒ Entre os primodoadores voluntários, 57,1% retornaram e iniciaram o segundo ciclo de doação, porém, o tempo para finalização do primeiro ciclo de doação entre eles foi extremamente grande, variando de 61 a 3084 dias. Após um ano decorrido da primeira doação, 40% deles haviam retornado para uma nova doação. Além disso, ao longo do seguimento, verificamos que 54% retornaram após dois anos, 59% retornaram após três anos, 62% após quatro anos e 64% após cinco anos. A partir do sexto ano, a propensão de retorno do doador é extremamente pequena indicando a ocorrência rara desse evento.
- ⇒ A partir de uma abordagem paramétrica com a construção de modelos de mistura, descrevemos o comportamento de retorno do doador com o intuito de verificar o perfil do candidato a doar mais associado ao retorno e ao não retorno para uma nova doação.
- ⇒ O ajuste de um modelo paramétrico assumindo uma distribuição Gama Generalizada Estendida (GGE) para o tempo entre a primeira e a segunda doação foi o modelo mais satisfatório entre os estudados, no qual estimou uma proporção de doadores improváveis de retornar de 20%.

-
- ⇒ Quando consideramos o perfil do doador como possíveis covariáveis associadas com esse tempo, observamos, na maioria delas, um efeito considerável apenas na parte logística. Isso significa que dependendo do perfil do doador, é esperada uma fração de não retorno maior.
- ⇒ Apesar do modelo ser considerado satisfatório em determinar o comportamento de retorno doador, ele se mostrou incapaz de discriminar perfis de doadores improváveis de retornar baseado na proporção p para outras amostras
- ⇒ Ao estimarmos uma proporção de doadores improváveis de retornar de 20% a partir do modelo assumindo uma distribuição GGE, não pudemos identificar qual período a partir do qual o retorno já não é mais observado, dado que este sugere ser além do período observado. Por outro lado, ao assumirmos uma distribuição Log-Normal para o tempo entre a primeira e a segunda doação (segundo modelo melhor ajustado), observamos que 34% dos doadores jamais retornarão, sendo este indicativo visto a partir de aproximadamente quatro anos após a primeira doação. Do ponto de vista prático, é viável considerar uma proporção de primodoadores que jamais retornarão de 34%.

CAPÍTULO VII

Perspectivas Futuras

Durante o período de mestrado, várias idéias surgiram e foram parcialmente realizadas diante dessa questão da doação de sangue.

1. Descrição do histórico de doação e de doadores em cada ano de seguimento.

O intuito é verificar a participação dos primodoadores e posteriormente dos primodoadores voluntários ao longo dos anos na doação de sangue. Além disso, há a possibilidade de classificar os doadores, de acordo com a proposta da ANVISA, ano a ano, porém, incorporando as opções: não retornou naquele ano e não retornou mais durante o período observado. Tais informações poderão simbolizar em números os esforços na captação e fidelização do doador ao longo do tempo.

2. Construção de um indicador de classificação do doador de sangue

A partir da construção de um indicador utilizando as informações: número de doações realizadas e número de anos de seguimento, propor uma classificação do doador de sangue. Esse indicador nos trará uma idéia do tempo de “vida útil” do mesmo e possivelmente do perfil do doador associado a um baixo e alto retorno.

3. Estudo do histórico de doação a partir do número de doações realizadas a partir de modelos lineares generalizados e modelos de mistura.

Através da abordagem dos Modelos Lineares Generalizados, introduzir modelos voltados à descrição do comportamento dos doadores de sangue e com isso poder estimar o número de doações por doador-ano (taxa de retorno) (SCREIBER et al., 2003). Além disso, há o interesse de observar o impacto da diminuição do número de doadores que não retornaram nessas taxas e consequentemente no número de bolsas de sangue a serem coletadas.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, H. J. **Aplicação da análise de sobrevivência em um problema de *Credit Scoring* e comparação com a regressão logística**. 2005. 116f. Dissertação (Mestrado em Estatística) – Departamento de Estatística, Universidade Federal de São Carlos, São Carlos, 2005.

BORGES, V. L.; MARTINEZ, E. Z.; BENDINI, M. H.; COSTA, M. A. G. F.; FERREIRA, S. C. L. Avaliação da fidedignidade de um instrumento voltado à satisfação do doador de sangue. **Revista Brasileira de Epidemiologia**, v. 8, n. 2, p. 177-186, 2005.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Relatório anual da área de Sangue e Hemoderivados**. Relatório de atividades de 1998. 1998. Disponível em: <<http://www.anvisa.gov.br/sangue/hemoterapia/producao.htm>>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Sistema de Legislação em Vigilância Sanitária**. Resolução RDC nº149 de 14 de agosto de 2001. Disponível em: <<http://www.anvisa.gov.br/legis/index.htm>>. Acesso em: 20 jan. 2007.

_____. Avança Brasil. Relatório Anual de Avaliação Prurianual 2002. **Avaliação dos Programas - Qualidade do Sangue**. 2002a. Disponível em: <<http://www.abrasil.gov.br/avalppa/RelAvalPPA2002/default.htm>>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Produção da Rede Hemoterápica**. 2002b. Disponível em: <<http://www.anvisa.gov.br/sangue/hemoterapia/producao.htm>>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Sistema de Legislação em Vigilância Sanitária**. Resolução RDC nº343 de 13 de dezembro de 2002c. Disponível em: <<http://www.anvisa.gov.br/legis/index.htm>>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Sistema de Legislação em Vigilância Sanitária**. Resolução RDC nº153 de 14 de junho de 2004a. Disponível em: <<http://www.anvisa.gov.br/legis/index.htm>>. Acesso em: 12 jun. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Notícias**. Brasil mobilizado para aumentar a doação de sangue. 2004b. Disponível em: <http://www.anvisa.gov.br/divulga/noticias/index1_2004.htm>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Perfil do Doador de Sangue Brasileiro**. 2004c. Disponível em: <<http://www.anvisa.gov.br/sangue/index.htm>>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Notícias**. Agência esclarece questionamento sobre doação de sangue. 2006. Disponível em: <<http://www.anvisa.gov.br/divulga/noticia/2006/310706.htm>>. Acesso em: 17 ago. 2007.

_____. Ministério da Saúde. **Programa Nacional de DST e AIDS - Portal Informativo sobre AIDS e outras doenças sexualmente transmissíveis**. Busca de notícias. Doação de sangue/Doador de sangue. S.d.a Disponível em: <<http://sistemas.aids.gov.br/imprensa/BuscaAvancada.asp?txtBusca=>>. Acesso em: 20 jan. 2007.

_____. Ministério da Saúde. Agência Nacional de Vigilância Sanitária (ANVISA). **Campanha Nacional de doação de sangue**. S.d.b Disponível em: <<http://www.anvisa.gov.br/sangue/campanha/index.htm>>. Acesso em: 20 jan. 2007.

COLOSIMO, E. A; GIOLO, S. R. **Análise de Sobrevida Aplicada**. São Paulo: Edgard Blucher, 2006.

CORBIERE, F; JOLY, P. A SAS macro for parametric e semiparametric mixture cure models. **Computer Methods and Programs in Biomedicine**, v. 85, p. 173-180, 2007.

COX, D. R. Regression models and life-tables (with discussion). **Journal of the Royal Statistic Society, B**, v. 74, p. 187-220, 1972.

FLEGEL, W. A.; BESENFELDER, W.; WAGNER, F. F. Predicting a donor's likelihood of donating within a preselected time interval. **Transfusion Medicine**, v. 10, n. 3, p. 181-192, 2000.

FRANKEL, P.; LONGMATE, J. Parametric models for accelerated and long-term survival: a comment on proportional hazards. **Statistics in Medicine**, v. 21, p. 3279-3289, 2002.

GLYNN, S. A.; KLEINMAN, S. H.; SCHREIBER, G. B.; ZUCH, T.; McCOMBS, S.; BETHEL, J.; GARRATTY, G.; WILLIAMS, A. E. Motivations to donate blood: demographic comparisons. **Transfusion**, v. 42, p. 216-225, 2002.

GLYNN, S. A.; BUSCH, M. P.; SCHREIBER, G. B.; MURPHY, E. L.; WRIGHT, D. J.; TU, Y.; KLEINMAN, S. H. Effect of a national disaster on blood supply and safety – The september 11 experience. **Journal American Medical Association**, v. 289, n. 17, p. 2246-2253, 2003.

GREENACRE, M. J. **Correspondence analysis in practice**. London: Academic Press, 1993.

GUERRA, C. C. C. Fim da doação remunerada de sangue no Brasil faz 25 anos. **Revista Brasileira de Hematologia e Hemoterapia**, v. 27, n. 1, p. 1-4, 2005.

HALPERIM, D.; BAETENS, J.; NEWMAN, B. The effect short-term, temporary deferral on future blood donation. **Transfusion**, v. 38, n. 2, p. 181-183, 1998.

HEMOCENTRO DE RIBEIRÃO PRETO - Centro Regional de Hemoterapia do HCFMRP - USP. Disponível em <<http://pegasus.fmrp.usp.br>>. Acesso em: 20 jan. 2007.

HOUGAARD, P. Fundamentals of survival data. **Biometrics**, v. 55, n. 1, p. 13-22, 1999.

JAMES, R. C.; MATTHEWS, D. E. The donation cycle: a framework for the measurement and analysis of blood donor return behaviour. **Vox Sanguinis**, v. 64, p. 37-42, 1993.

JAMES, R. C.; MATTHEWS, D. E. Analysis of blood donor return behaviour using survival regression methods. **Transfusion Medicine**, v. 6, n. 1, p. 21-30, 1996.

JUNQUEIRA, P. C.; ROSENBLIT, J.; HAMERSCHLAK, N. História da Hemoterapia no Brasil. **Revista Brasileira de Hematologia e Hemoterapia**, v. 27, n. 3, p. 201-207, 2005.

KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American Statistics Association**, v. 53, p. 457-481, 1958.

- LEE, E. T.; WANG, J. W. **Statistical Methods for Survival Data Analysis**, John Wiley & Sons, 2003.
- LOUZADA-NETO, F.; MAZUCHELI, J.; ACHCAR, J. A. **Análise de Sobrevida e Confiabilidade**. Lima: IMCA, 2002. (Monografias del IMCA, n.30).
- LUDWIG, S. T.; RODRIGUES, A. C. M. Doação de Sangue: uma visão de marketing. **Cadernos de Saúde Pública**, v. 21, n. 3, p. 932-939, 2005.
- MALLER, R. A.; ZHOU, X. **Survival analysis with long-term survivors**. John Wiley, New York, 1996.
- MAZUCHELI, J.; LOUZADA-NETO, F.; ACHCAR, J. A. Lifetime models with noconstant shape parameters. **Revstat**, v. 1, p. 25-39, 2003.
- MELNYK, S. A.; PAGELL, M.; JORAE, G.; SHARPE, A. S. Applying survival analysis to operations management: analyzing the differences in donor classes in the blood donation process. **Journal of Operations Management**, v. 13, n. 4, p. 339-356, 1995.
- MISJE, A. H.; BOSNES, V.; GASDAL, O.; HEIER, H. E. Motivation, recruitment and retention of voluntary non-remunerated blood donors: a survey-based questionnaire study. **Vox Sanguinis**, v. 89, p. 236-244, 2005.
- MONTOYA, A. F.; CASTILLO, J. D. L.; BERRIO, A. L.; FERNANDEZ, A. R. Atitudes, creencias y motivaciones em donantes de sangre. **Sangre**, v. 41, n. 6, p. 427-440, 1996.
- MOURA, A. S.; MOREIRA, C. T.; MACHADO, C. A.; NETO, J. A. V.; MACHADO, M. F. A. S. Doador de Sangue habitual e fidelizado: fatores motivacionais de adesão ao programa. **Revista Brasileira em Promoção da Saúde**, v. 19, n. 2, p. 61-67, 2006.
- OWNBY, H. E.; WATANABE, Y. Tu; NASS, C. C. Analysis of donor return behavior. **Transfusion**, v. 39, n. 10, p. 1128-1135, 1999.
- PASSOS, A. D. C.; FARINA, R. M. A. Conhecimentos, atitudes e práticas em relação à doação sanguínea entre acadêmicos de medicina. **Revista de Medicina de Ribeirão Preto**, v. 27, n. 3/4, p. 380-388, 2004.

PENG, Y; DEAR, K. B. G.; DENHAM, J. W. A generalized F mixture model for cure rate estimation. **Statistics in Medicine**, v. 17, n. 8, p. 813-830, 1998.

PENG, Y; DEAR, K. B. G. A nonparametric mixture model for cure rate estimation. **Biometrics**, v. 56, n. 1, p. 237-243, 2000.

PENG, Y; CARRIERE, K. C. An empirical comparison of parametric and semiparametric cure models. **Biometrical Journal**, v. 44, p. 1002-1014, 2002.

PILLIAVIN, J. A. Temporary deferral and donor return. **Transfusion**, v. 27, n. 2, p. 199-200, 1987.

QUIDIM, I. L. **Análise de sobrevivência com fração de fidelizados: uma aplicação na área de Marketing**. 2005, 78f. Dissertação (Mestrado em Estatística) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2005.

RIZZATO, F. B. **Modelos de regressão log-gama generalizado com fração de cura**. 2006, 73f. Dissertação (Mestrado em Agronomia) – Escola Superior de Agricultura “Luiz de Queirós”, Universidade de São Paulo, São Paulo, 2006.

SCHEMPER, M. Cox analysis of survival data with non-proportional hazard functions. **The Statistician**, v. 41, n. 4, p. 455-465, 1992.

SCHREIBER, G. B.; SANCHES, A. M.; GLYNN, S. A.; WRIGHT, D. J. Increasing blood availability by changing donation patterns. **Transfusion**, v. 43, p. 591-597, 2003.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**. v. 6, n. 2, p. 461-464, 1978.

SIMON, T. Where have all the donors gone? A personal reflection on the crisis in America's volunteer blood program. **Transfusion**, v. 43, n. 2, p. 273-279, 2003.

SPOSTO, R. Cure model analysis in cancer: an application to data from the Children's Cancer Group. **Statistics in Medicine**, v. 21, n. 2, p. 293-312, 2002.

SUAREZ, I. M. B.; MONTOYA, A. F.; FERNANDEZ, A. R.; BERRIO, A. L.; PEÑUELA, M. C. How regular blood donors explain their behavior. **Transfusion**, v. 44, n. 10, p. 1441-1446, 2004.

SUGIURA, N. Further analysis of the data by Akaike's information criterion and the finite correction. **Communications in Statistics – Theory and Methods**, v. 7, n. 1, p. 13-26, 1978.

THOMAS, L. C.; EDELMAN, D. B.; CROOK, J. N. **Credit Scoring and its Applications**. SIAM, 2002.

VICARI, V. M. **Análise dos fatores determinantes para a doação de sangue: o caso do Hemocentro de Ribeirão Preto**. 2003, 139f. Dissertação (Mestrado em Engenharia de Produção) – Departamento de Engenharia de Produção, Universidade de São Paulo, São Carlos, 2003.

YAMAGUCHI, K. Accelerated failure-time regression models with a regression model of survival fraction: An application to the analysis of permanent employment in Japan. **Journal of the American Statistical Association**, v. 87, p. 284-292, 1992.

WHYTE, G. Quantitating donor behaviour to model the effect of changes in donor management on sufficiency in the blood service. **Vox Sanguinis**, v. 76, n. 4, p. 209-215, 1999.

ZILMER, E. A.; GLIDDEN, R. A.; HONAKER, L. M.; MEYER J. D. Mood states in the volunteer blood donor. **Transfusion**, v. 29, n. 1, p. 27-30, 1989.

ZWEIG, M. H.; CAMPBELL, G. Receiver operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. **Clinical Chemistry**, v. 39, n. 4, p. 561-577, 1993.

ANEXOS

ANEXO A – Mapa da Região de abrangência de Ribeirão Preto

ANEXO B – Autorização

DECLARAÇÃO

Declaro estar ciente que informações relativas às características dos doadores de sangue do Hemocentro de Ribeirão Preto serão utilizadas para a realização da pesquisa *“INTRODUÇÃO DE UM MODELO DE REGRESSÃO BAYESIANO DE EVENTOS RECORRENTES APLICADO AO ESTUDO DO COMPORTAMENTO DE RETORNO DE DOADORES DE SANGUE DO HEMOCENTRO DE RIBEIRÃO PRETO, SP”*, realizada pela pesquisadora Adriana de Fátima Lourençon, Bacharel em Estatística, aluna do Programa de Pós-Graduação em Saúde da Comunidade da Faculdade de Medicina de Ribeirão Preto – FMRP/USP.

Ribeirão Preto, 21 de março de 2005.

Prof. Dr. Dimas Tadeu Covas
Diretor do Hemocentro de Ribeirão Preto

ANEXO C – Aprovação pelo Comitê de Ética em Pesquisa do HCFMRP - USP

**HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO**

CAMPUS UNIVERSITÁRIO – MONTE ALEGRE
FONE: 602-1000 - FAX (016) 633-1144

Ribeirão Preto, 14 de abril de 2005

Ofício nº 991/2005
CEP/CDGC

Prezada Senhora:

O trabalho intitulado **“INTRODUÇÃO DE UM MODELO DE REGRESSÃO BAYESIANO DE EVENTOS RECORRENTES APLICADO AO ESTUDO DO COMPORTAMENTO DE RETORNO DE DOADORES DE SANGUE”**, foi analisado pelo Comitê de Ética em Pesquisa, em sua 199ª Reunião Ordinária realizada em 11/04/2005, e enquadrado na categoria: **APROVADO, bem como o Termo de Consentimento Livre e Esclarecido**, de acordo com o Processo HCRP nº 3907/2005.

Aproveito a oportunidade para apresentar a Vossa Senhoria protestos de estima e consideração.

PROFª DRª CLARISSE DULCE G. CARVALHEIRO
Vice-Coordenadora do Comitê de Ética
em Pesquisa do HCRP e da FMRP-USP

Ilustríssima Senhora
ADRIANA DE FÁTIMA LOURENÇON
PROF. DR. EDSON ZANGIACOMI MARTINEZ (Orientador)
Centro de Métodos Quantitativos – CEMEQ – FMRP-USP
Em mãos

ANEXO D – Funções de Verossimilhanças dos Modelos Paramétricos

Modelo Paramétrico:

Seja $T_1, \dots, T_n$ uma variável aleatória representando o tempo entre a primeira e a segunda doação de n doadores de sangue e que associado a ela exista um vetor de k covariáveis, $\mathbf{x}$, atribuído ao i -ésimo doador de sangue, e uma variável indicadora δ_i , onde $\delta_i = 1$ se for observado um retorno do i -ésimo doador de sangue e $\delta_i = 0$ se até o fim do seguimento esse evento não foi observado. A maximização da função de verossimilhança, de (3.34), utilizando a relação (3.4), pode ser escrita como,

$$L(\hat{\lambda}, \hat{\gamma} / \mathbf{x}) = \sum_{i=1}^n \left\{ \delta_i \ln [f(t_i / \mathbf{x}_i)] + (1 - \delta_i) \ln [S(t_i / \mathbf{x}_i)] \right\} \quad (\text{A1})$$

$$= \sum_{i=1}^n \left\{ \delta_i \ln [h(t_i / \mathbf{x}_i) S(t_i / \mathbf{x}_i)] + (1 - \delta_i) \ln [S(t_i / \mathbf{x}_i)] \right\} \quad (\text{A2})$$

$$= \sum_{i=1}^n \left\{ \delta_i [\ln h(t_i / \mathbf{x}_i) + \ln S(t_i / \mathbf{x}_i)] + (1 - \delta_i) [\ln S(t_i / \mathbf{x}_i)] \right\} \quad (\text{A3})$$

$$= \sum_{i=1}^n \left\{ \delta_i \ln h(t_i / \mathbf{x}_i) + \delta_i \ln S(t_i / \mathbf{x}_i) + \ln S(t_i / \mathbf{x}_i) - \delta_i \ln S(t_i / \mathbf{x}_i) \right\} \quad (\text{A4})$$

$$= \sum_{i=1}^n \left\{ \delta_i \ln h(t_i / \mathbf{x}_i) + \ln S(t_i / \mathbf{x}_i) \right\} \quad (\text{A5})$$

As respectivas funções de verossimilhanças considerando as distribuições de probabilidade: Weibull, Log-Normal, Log-Logística e Gama Generalizada Extendida são descritas abaixo.

Distribuição Weibull

Considere um modelo Weibull com função densidade de probabilidade dada em (3.14) e função de sobrevivência dada em (3.15). Assumindo que o parâmetro de locação é afetado por um vetor de k covariáveis através da relação (3.30), a função de verossimilhança, partindo da equação (A5), é dada por,

$$\begin{aligned}
 L(\hat{\alpha}, \hat{\gamma} / \mathbf{x}) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[\lambda(\mathbf{x}_i) \gamma (\lambda(\mathbf{x}_i) t_i)^{\gamma-1} \right] + \ln \left[e^{-(\lambda(\mathbf{x}_i) t_i)^{\gamma}} \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \ln \left[e^{x_i \alpha} \gamma (e^{x_i \alpha} t_i)^{\gamma-1} \right] + \ln \left[e^{-(e^{x_i \alpha} t_i)^{\gamma}} \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\mathbf{x}_i \alpha + \ln \gamma + (\gamma-1) \ln(e^{x_i \alpha} t_i) \right] - (e^{x_i \alpha} t_i)^{\gamma} \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln \gamma + \gamma \mathbf{x}_i \alpha + \gamma \ln t_i - \ln t_i \right] - (e^{x_i \alpha} t_i)^{\gamma} \right\}
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$ e $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$.

Distribuição Log-Normal

Considere um modelo Log-Normal com função densidade de probabilidade dada em (3.20) e função de sobrevivência dada em (3.21). Assumindo que o parâmetro de localização é afetado por um vetor de k covariáveis através da relação (3.30), a função de verossimilhança, partindo da equação (A1), é dada por,

$$\begin{aligned}
 L(\hat{\alpha}, \hat{\sigma} / \mathbf{x}) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[\frac{1}{\sqrt{2\pi}\sigma t_i} \exp \left[-\frac{1}{2\sigma^2} \left(\ln(\lambda(\mathbf{x}_i)t_i) \right)^2 \right] \right] + (1-\delta_i) \ln \left[1 - \Phi \left(\ln \frac{\lambda(\mathbf{x}_i)t_i}{\sigma} \right) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[-\ln(\sqrt{2\pi}\sigma t_i) - \left[\frac{1}{2\sigma^2} \left(\ln(e^{\mathbf{x}_i \alpha} t_i) \right)^2 \right] \right] + (1-\delta_i) \ln \left[1 - \Phi \left(\ln \frac{e^{\mathbf{x}_i \alpha} t_i}{\sigma} \right) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[-\ln(\sqrt{2\pi}\sigma t_i) - \frac{1}{2\sigma^2} (\mathbf{x}_i \alpha + \ln t_i)^2 \right] + (1-\delta_i) \ln \left[1 - \Phi \left(\ln \frac{e^{\mathbf{x}_i \alpha} t_i}{\sigma} \right) \right] \right\}
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$ e $\Phi(\cdot)$ é a função distribuição acumulada da distribuição normal.

Distribuição Log-Logística

Considere um modelo Log-Logística com função densidade de probabilidade dada em (3.22) e função de sobrevivência dada em (3.23). Assumindo que o parâmetro de locação é afetado por um vetor de k covariáveis através da relação (3.30), a função de verossimilhança, partindo da equação (A5), é dada por,

$$\begin{aligned}
 L(\hat{\boldsymbol{\alpha}}, \hat{\gamma} / \mathbf{x}) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[\frac{\lambda(\mathbf{x}_i) \gamma t_i^{\gamma-1}}{(1 + \lambda(\mathbf{x}_i) t_i^{\gamma})} \right] + \ln \left[\frac{1}{(1 + \lambda(\mathbf{x}_i) t_i^{\gamma})} \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \ln \left[\frac{e^{x_i \boldsymbol{\alpha}} \gamma t_i^{\gamma-1}}{(1 + e^{x_i \boldsymbol{\alpha}} t_i^{\gamma})} \right] + \ln \left[\frac{1}{(1 + e^{x_i \boldsymbol{\alpha}} t_i^{\gamma})} \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(e^{x_i \boldsymbol{\alpha}} \gamma t_i^{\gamma-1}) - \ln(1 + e^{x_i \boldsymbol{\alpha}} t_i^{\gamma}) \right] - \ln(1 + e^{x_i \boldsymbol{\alpha}} t_i^{\gamma}) \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\mathbf{x}_i \boldsymbol{\alpha} + \ln \gamma + (\gamma - 1) \ln t_i - \ln(1 + e^{x_i \boldsymbol{\alpha}} t_i^{\gamma}) \right] - \ln(1 + e^{x_i \boldsymbol{\alpha}} t_i^{\gamma}) \right\}
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$ e $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$.

Distribuição Gama Generalizada Extendida (GGE)

Considere um modelo GGE com função densidade de probabilidade dada em (3.28) e função de sobrevivência dada em (3.29). Assumindo que o parâmetro de locação é afetado por um vetor de k covariáveis através da relação (3.30), a função de verossimilhança quando $\gamma < 0$, partindo da equação (A1), é dada por,

$$L(\hat{\alpha}, \hat{\gamma} / \mathbf{x}) = \sum_{i=1}^n \left\{ \delta_i \ln \left[\frac{|\gamma| \alpha^\alpha \lambda(\mathbf{x}_i)^{\alpha\gamma}}{\Gamma(\alpha)} t_i^{\alpha\gamma-1} \exp \left[-\alpha (\lambda(\mathbf{x}_i) t_i)^\gamma \right] \right] + (1 - \delta_i) \ln \left[I \left(\alpha (\lambda(\mathbf{x}_i) t_i)^\gamma, \alpha \right) \right] \right\}$$

$$L(\hat{\alpha}, \hat{\gamma} / \mathbf{x}) = \sum_{i=1}^n \left\{ \delta_i \ln \left[\frac{|\gamma| \alpha^\alpha (e^{x_i \alpha})^{\alpha\gamma}}{\Gamma(\alpha)} t_i^{\alpha\gamma-1} \exp \left[-\alpha (e^{x_i \alpha} t_i)^\gamma \right] \right] + (1 - \delta_i) \ln \left[I \left(\alpha (e^{x_i \alpha} t_i)^\gamma, \alpha \right) \right] \right\}$$

$$L(\hat{\alpha}, \hat{\gamma} / \mathbf{x}) = \sum_{i=1}^n \left\{ \delta_i \left[\ln \left(\frac{|\gamma| \alpha^\alpha (e^{x_i \alpha})^{\alpha\gamma}}{\Gamma(\alpha)} \right) + (\alpha\gamma - 1) \ln t_i - \alpha (e^{x_i \alpha} t_i)^\gamma \right] + (1 - \delta_i) \ln \left[I \left(\alpha (e^{x_i \alpha} t_i)^\gamma, \alpha \right) \right] \right\}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$ e $I(\cdot)$ é a função gama incompleta.

ANEXO E – Funções de Verossimilhanças dos Modelos Paramétricos Longa-Duração

Modelo Paramétrico Longa-Duração (MLD):

Seja $T_1, \dots, T_n$ uma variável aleatória representando os tempos entre a primeira e a segunda doação de n doadores de sangue e que associado a ela exista um vetor de k covariáveis, $\mathbf{x}$, e m covariáveis, $\mathbf{z}$, atribuído ao i -ésimo doador, e uma variável indicadora δ_i , onde $\delta_i = 1$ se for observado um retorno do i -ésimo doador de sangue na amostra e $\delta_i = 0$ se não for observado esse evento.

Considere também um modelo paramétrico longa-duração com função densidade de probabilidade dada em (3.42) e função de sobrevivência dada em (3.41). A maximização da função de verossimilhança, de (3.45), pode ser escrita como,

$$L(\hat{\alpha}, \hat{\beta}, \gamma / \mathbf{x}, \mathbf{z}) = \sum_{i=1}^n \left\{ \delta_i \ln [f(t_i / \mathbf{x}_i, \mathbf{z}_i)] + (1 - \delta_i) \ln [S(t_i / \mathbf{x}_i, \mathbf{z}_i)] \right\} \quad (\text{A6})$$

$$= \sum_{i=1}^n \left\{ \delta_i \ln [(1 - p(\mathbf{z}_i)) f_0(t_i / \mathbf{x}_i, \mathbf{z}_i)] + (1 - \delta_i) \ln [p(\mathbf{z}_i) + (1 - p(\mathbf{z}_i)) S_0(t_i / \mathbf{x}_i, \mathbf{z}_i)] \right\} \quad (\text{A7})$$

As respectivas funções de verossimilhanças considerando as distribuições de probabilidade: Weibull, Log-Normal, Log-Logística e Gama Generalizada Extendida são descritas abaixo.

MLD: Distribuição Weibull

Considere um modelo longa-duração Weibull com função densidade de probabilidade dada em (3.14) e função de sobrevivência dada em (3.15). Vamos assumir que o parâmetro de locação é afetado por um vetor de k covariáveis, através da relação (3.30) e que a proporção de doadores improváveis de retornar, representado por, p , é afetada por um vetor de m covariáveis, através da relação (3.43). A função de verossimilhança, partindo da equação (A7), é dada por,

$$\begin{aligned}
 L(\hat{\alpha}, \hat{\beta}, \gamma) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[(1 - p(z_i)) \lambda(\mathbf{x}_i) \gamma (\lambda(\mathbf{x}_i) t_i)^{\gamma-1} e^{-(\lambda(\mathbf{x}_i) t_i)^\gamma} \right] \right\} + \\
 &\quad \sum_{i=1}^n (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) e^{-(\lambda(\mathbf{x}_i) t_i)^\gamma} \right] \\
 &= \sum_{i=1}^n \left\{ \delta_i \ln \left[(1 - p(z_i)) e^{\mathbf{x}_i \alpha} \gamma (e^{\mathbf{x}_i \alpha} t_i)^{\gamma-1} e^{-(e^{\mathbf{x}_i \alpha} t_i)^\gamma} \right] \right\} + \\
 &\quad \sum_{i=1}^n (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) e^{-(e^{\mathbf{x}_i \alpha} t_i)^\gamma} \right] \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1 - p(z_i)) + \mathbf{x}_i \alpha + \ln \gamma + (\gamma - 1) \ln(e^{\mathbf{x}_i \alpha} t_i) - (e^{\mathbf{x}_i \alpha} t_i)^\gamma \right] \right\} + \\
 &\quad \sum_{i=1}^n (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) e^{-(e^{\mathbf{x}_i \alpha} t_i)^\gamma} \right] \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1 - p(z_i)) + \ln \gamma + \gamma \mathbf{x}_i \alpha + \gamma \ln t_i - \ln t_i - (e^{\mathbf{x}_i \alpha} t_i)^\gamma \right] \right\} + \\
 &\quad \sum_{i=1}^n (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) e^{-(e^{\mathbf{x}_i \alpha} t_i)^\gamma} \right]
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, $\mathbf{z}_i = (1, z_{i1}, z_{i2}, \dots, z_{ik})^T$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$,

$\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_m)^T$ e $p(z_i) = [1 + e^{z_i \beta}]^{-1}$.

MLD: Distribuição Log-Normal

Considere um modelo longa-duração Log-Normal com função densidade de probabilidade dada em (3.20) e função de sobrevivência dada em (3.21). Vamos assumir que o parâmetro de locação é afetado por um vetor de k covariáveis, através da relação (3.30) e que a proporção de doadores improváveis de retornar, representado por, p , é afetada por um vetor de m covariáveis, através da relação (3.43). A função de verossimilhança, partindo da equação (A7), é dada por,

$$\begin{aligned}
 L(\hat{\alpha}, \hat{\beta}, \gamma / \mathbf{x}, \mathbf{z}) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[(1 - p(z_i)) \frac{1}{\sqrt{2\pi}\sigma t_i} \exp \left[-\frac{1}{2\sigma^2} \left(\ln(\lambda(\mathbf{x}_i)t_i) \right)^2 \right] \right] \right\} + \\
 &\quad \sum_{i=1}^n \left\{ (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) 1 - \Phi \left(\ln \frac{\lambda(\mathbf{x}_i)t_i}{\sigma} \right) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1 - p(z_i)) - \ln(\sqrt{2\pi}\sigma t_i) - \frac{1}{2\sigma^2} \left(\ln(e^{x_i \alpha} t_i) \right)^2 \right] \right\} + \\
 &\quad \sum_{i=1}^n \left\{ (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) 1 - \Phi \left(\ln \frac{e^{x_i \alpha} t_i}{\sigma} \right) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1 - p(z_i)) - \ln(\sqrt{2\pi}\sigma t_i) - \frac{1}{2\sigma^2} (\mathbf{x}_i \boldsymbol{\alpha} + \ln t_i)^2 \right] \right\} + \\
 &\quad \sum_{i=1}^n \left\{ (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) 1 - \Phi \left(\ln \frac{e^{x_i \alpha} t_i}{\sigma} \right) \right] \right\}
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, $\mathbf{z}_i = (1, z_{i1}, z_{i2}, \dots, z_{ik})^T$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$,

$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_m)^T$, $p(z_i) = [1 + e^{z_i \boldsymbol{\beta}}]^{-1}$ e $\Phi(\cdot)$ é a função distribuição acumulada da distribuição normal.

.

MLD: Distribuição Log-Logística

Considere um modelo longa-duração Log-Logística com função densidade de probabilidade dada em (3.22) e função de sobrevivência dada em (3.23). Vamos assumir que o parâmetro de locação é afetado por um vetor de k covariáveis, através da relação (3.30) e que a proporção de doadores improváveis de retornar, representado por, p , é afetada por um vetor de m covariáveis, através da relação (3.43). A função de verossimilhança, partindo da equação (A7), é dada por,

$$\begin{aligned}
 L(\hat{\alpha}, \hat{\beta}, \gamma) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[(1 - p(z_i)) \frac{\lambda(\mathbf{x}_i) \gamma t_i^{\gamma-1}}{(1 + \lambda(\mathbf{x}_i) t_i^\gamma)^2} \right] + (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) \frac{1}{(1 + \lambda(\mathbf{x}_i) t_i^\gamma)} \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1 - p(z_i)) + \ln(e^{\mathbf{x}_i \alpha} \gamma t_i^{\gamma-1}) - \ln(1 + e^{\mathbf{x}_i \alpha} t_i^\gamma)^2 \right] \right\} + \\
 &\quad \sum_{i=1}^n \left\{ (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) \frac{1}{(1 + \lambda(\mathbf{x}_i) t_i^\gamma)} \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1 - p(z_i)) + \mathbf{x}_i \alpha + \ln \gamma + (\gamma - 1) \ln t_i - \ln(1 + e^{\mathbf{x}_i \alpha} t_i^\gamma)^2 \right] \right\} + \\
 &\quad \sum_{i=1}^n \left\{ (1 - \delta_i) \ln \left[p(z_i) + (1 - p(z_i)) \frac{1}{(1 + e^{\mathbf{x}_i \alpha} t_i^\gamma)} \right] \right\}
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, $\mathbf{z}_i = (1, z_{i1}, z_{i2}, \dots, z_{ik})^T$, $\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$,

$\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_m)^T$ e $p(z_i) = [1 + e^{z_i \beta}]^{-1}$.

MLD: Distribuição Gama Generalizada Extendida (GEE)

Considere um modelo GGE com função densidade de probabilidade dada em (3.28) e função de sobrevivência dada em (3.29). Vamos assumir que o parâmetro de locação é afetado por um vetor de k covariáveis, através da relação (3.30) e que a proporção de doadores improváveis de retornar, representado por, p , é afetada por um vetor de m covariáveis, através da relação (3.43). A função de verossimilhança, partindo da equação (A7), é dada por,

$$\begin{aligned}
 L(\hat{\alpha}, \hat{\gamma} / \mathbf{x}) &= \sum_{i=1}^n \left\{ \delta_i \ln \left[(1-p(z_i)) \frac{/\gamma / \alpha^\alpha \lambda(\mathbf{x}_i)^{\alpha\gamma}}{\Gamma(\alpha)} t_i^{\alpha\gamma-1} \exp \left[-\alpha (\lambda(\mathbf{x}_i) t_i)^\gamma \right] \right] + \right. \\
 &\quad \left. (1-\delta_i) \ln \left[p(z_i) + (1-p(z_i)) I \left(\alpha (\lambda(\mathbf{x}_i) t_i)^\gamma, \alpha \right) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \ln \left[(1-p(z_i)) \frac{/\gamma / \alpha^\alpha (e^{x_i \alpha})^{\alpha\gamma}}{\Gamma(\alpha)} t_i^{\alpha\gamma-1} \exp \left[-\alpha (e^{x_i \alpha} t_i)^\gamma \right] \right] + \right. \\
 &\quad \left. (1-\delta_i) \ln \left[p(z_i) + (1-p(z_i)) I \left(\alpha (e^{x_i \alpha} t_i)^\gamma, \alpha \right) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \delta_i \left[\ln(1-p(z_i)) + \ln \left(\frac{/\gamma / \alpha^\alpha (e^{x_i \alpha})^{\alpha\gamma}}{\Gamma(\alpha)} \right) + (\alpha\gamma-1) \ln t_i - \alpha (e^{x_i \alpha} t_i)^\gamma \right] + \right. \\
 &\quad \left. (1-\delta_i) \ln \left[p(z_i) + (1-p(z_i)) I \left(\alpha (e^{x_i \alpha} t_i)^\gamma, \alpha \right) \right] \right\}
 \end{aligned}$$

onde $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, $\mathbf{z}_i = (1, z_{i1}, z_{i2}, \dots, z_{ik})^T$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)^T$,

$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_m)^T$, $p(z_i) = [1 + e^{z_i \boldsymbol{\beta}}]^{-1}$ e $I(\cdot)$ é a função gama incompleta.

ANEXO F – Resultados obtidos com as respectivas Análises de Correspondências

Figura F1. Visualização das associações entre as categorias das variáveis relacionadas com o perfil sócio-demográfico do primodoador voluntário e a sua condição até o fim do seguimento (retornou ou não para uma nova doação) através do mapa gerado pela análise de correspondência. Nos eixos dos gráficos, os valores em parênteses referem-se à inércia relativa à dimensão em questão, ou seja, a variação total dos dados decomposta em cada dimensão (ou eixo) do gráfico informando a proporção da variação total que aquele eixo explica.

LEGENDA:

Sexo:	Faixa Etária:	Estado Civil:	Tipo Sanguíneo:	Unidade de Coleta:
Mas: Masculino	F1: 18 a 19	Cas: Casado	A: A	CE: Coleta Externa
Fem: Feminino	F2: 20 a 29	Div: Divorciado	AB: AB	UP: Unidades Pequenas
	F3: 30 a 39	Out: Outros	B: B	UG: Unidades Grandes
	F4: 40 a 49	Sol: Solteiro	O: O	
	F5: 50 a 59	Viu: Viuvo		
	F6: 60 ou +			
Cor da Pele:			Fator Rh:	Situação:
Ama: Amarela			RhN: Negativo	Ret: Retornou
Bra: Branca			RhP: Positivo	N Ret: Não retorno
Neg: Negra

Figura F2. Visualização das associações entre as categorias das variáveis relacionadas com o perfil sócio-demográfico do primodoador voluntário e a sua condição até o fim do seguimento (retornou ou não para uma nova doação) através do mapa gerado pela análise de correspondência – Inclusão da variável nível de escolaridade. Nos eixos dos gráficos, os valores em parênteses referem-se à inércia relativa à dimensão em questão, ou seja, a variação total dos dados decomposta em cada dimensão (ou eixo) do gráfico informando a proporção da variação total que aquele eixo explica.

LEGENDA:

Sexo:	Faixa Etária:	Estado Civil:	Tipo Sanguíneo:	Escolaridade:	Unidade de Coleta:
Mas: Masculino	F1: 18 a 19	Cas: Casado	A: A	E1: Analfabeto	CE: Coleta Externa
Fem: Feminino	F2: 20 a 29	Div: Divorciado	AB: AB	E2: 1º Grau Incomp.	UP: Unidades Pequenas
	F3: 30 a 39	Out: Outros	B: B	E3: 1º Grau Comp./ 2º Grau Incomp.	UG: Unidades Grandes
	F4: 40 a 49	Sol: Solteiro	O: O	E4: 2º Grau Comp./ 3º Grau Incomp.	
Cor da Pele:	F5: 50 a 59	Viu: Viuvo		E5: 3º Grau Comp.	Situação:
Ama: Amarela	F6: 60 ou +		Fator Rh:	E6: Pós Graduação	Ret: Retornou
Bra: Branca			RhN: Negativo		N Ret: Não retornou
Neg: Negra			RhP: Positivo		

ANEXO G – Outros Resultados

Figura G1. Função de sobrevivência representando a propensão à não retornar dos primodoadores voluntários segundo aprovação na etapa de triagem e tipo de doação (gráfico superior) e respectiva função risco instantâneo à retornar (gráfico inferior). Notação: AP – doador aprovado e RT – doador reprovado temporariamente.

ANEXO H – Outros Resultados

Figura G2. Função de sobrevivência representando a propensão à não retornar dos primodoadores voluntários segundo ano da primeira doação (gráfico superior) e respectiva função risco instantâneo à retornar (gráfico inferior).

ANEXO I – Glossário

Ciclos de doações: é definido como uma série de quatro eventos subseqüentes: 1) inicia a partir da primeira doação, 2) depois há um espaço de tempo mínimo necessário para que o organismo se recupere, 3) seguido de um intervalo de tempo que o doador escolhe para doar novamente e 4) finaliza com uma subseqüente doação.

Coleta Externa: quando a doação é realizada numa unidade móvel de coleta de sangue.

Coleta Interna: quando a doação é realizada num posto de coleta, unidade, núcleo ou hemocentro.

Doação autóloga: é aquela doação advinda do indivíduo que doa para si próprio.

Doação de reposição: é aquela doação advinda do indivíduo que doa para atender a necessidade de um paciente, inclusive as doações dirigidas, vinculadas ou específicas.

Doação específica: o mesmo que doação de reposição.

Doação espontânea: é aquela doação advinda do indivíduo que procura o serviço de hemoterapia espontaneamente. Mesmo que doação voluntária.

Doação fracionada: a doação do sangue é total, sendo fracionada posteriormente nos diversos hemocomponentes do sangue para o seu melhor aproveitamento.

Doação por aféreses: significa um procedimento com separação de algum hemocomponente do sangue (plaquetas, hemáceas, plasma, granulócitos e linfócitos) através de um equipamento.

Doação Vinculada: o mesmo que doação de reposição.

Doação Voluntária: o mesmo que doação espontânea.

Doador de repetição: é aquele indivíduo que doa pelo menos uma vez a cada treze meses, considerando a data de sua última doação no mesmo serviço de hemoterapia.

Doador esporádico: é aquele indivíduo que doa eventualmente num intervalo superior a treze meses.

Doador fidelizado: o mesmo que doador regular.

Doador raro: denotado como a presença de um anticorpo ou antígeno não usual no soro do doador.

Doador regular: é aquele indivíduo que doa regularmente naquela unidade de hemoterapia. O mesmo que doador de repetição, habitual, fidelizado.

Doador inapto temporariamente: doador de sangue impossibilitado temporariamente de doar por não passar em alguma etapa do processo de triagem.

Doador inapto definitivamente: quando o sangue de um determinado indivíduo é considerado inutilizado por alguma razão, entre elas o exame laboratorial positivo para doenças transmissíveis.

Empírico: refere-se ao que é derivado de experimento ou de observação da realidade.

Exame extra rotina: etapa no qual o doador passa por todo o processo de triagem que antecede a doação mais não doa efetivamente uma bolsa de sangue, apenas colhe amostras para exames e aguarda um momento específico para doar, geralmente por convocação do serviço de hemoterapia.

Fator Rh: sistema no qual proporciona a produção do anticorpos por indução do antígeno. Existem seis tipos comuns de antígenos Rh, designados por C, D, E, c, d e e. Indivíduos cujas hemácias não se aglutinam na presença do soro anti-Rh, é portador de Rh negativo (c, d e e), caso contrário, são portadores de Rh positivo (C, D e E). Os doadores portadores de Rh negativo, embora menos prevalentes na população, podem doar para seus respectivos tipos sanguíneos portadores de Rh positivo. Por isso, são os tipos mais utilizados, principalmente em situações emergenciais.

Fidelização: ato de fidelizar, tornar fiel, permanente. Utilizado no sentido de se ter doadores de sangue que doam várias vezes;

Hemocentro: serviço de hemoterapia de alta complexidade de natureza pública, responsável pela diretriz da política de sangue em um estado ou numa determinada região de um estado.

Hemocomponentes: são os componentes do sangue. São eles: as células (hemácias, plaquetas, granulócitos) e a parte líquida, o plasma.

Hemoderivados: são as proteínas derivadas do plasma, obtidas por meio de processo industrial;

Hemoterapia: é a terapia por meio de hemocomponentes e hemoderivados;

Histórico de doação: número de ciclos de doação finalizados (número de doações realizadas) anteriormente.

Primodoador: indivíduo que doa pela primeira vez num determinado serviço de hemoterapia.

Sistema ABO: é um dos sistema de grupos sanguíneos mais importantes por ser o mais antigênico, ou seja, por ter maior capacidade de provocar a produção de anticorpos seguido pelo sistema Rh. Fazem parte desse sistema os três genes: A, B e O podendo qualquer um dos três ocupar o loco ABO em cada elemento do par de cromossomos responsáveis por este sistema. Ele é composto por quatro tipos: A, AB, B e O.

Triagem: Após a pré triagem, onde são verificadas a pressão arterial, pulso, temperatura corporal, peso, altura e hemoglobina (diagnostico ou não de anemia), o doador é entrevistado em uma sala em condições de privacidade por um profissional da área de saúde (médico ou enfermeiro). Posteriormente é realizada a coleta do sangue e servido um lanche.