

Universidade Federal do Paraná
Setor de Ciências Exatas
Departamento de Estatística
Programa de Especialização em *Data Science* e *Big Data*

Dayane Gonçalves Hermam

**Estudo de desenvolvimento de anomalias
congênitas pelo sistema do SUS**

**Curitiba
2022**

Dayane Gonçalves Hermam

Estudo de desenvolvimento de anomalias congênitas pelo sistema do SUS

Monografia apresentada ao Programa de Especialização em Data Science e Big Data da Universidade Federal do Paraná como requisito parcial para a obtenção do grau de especialista.

Orientador: José Luiz Padilha da Silva

Curitiba
2022

Estudo de desenvolvimento de anomalias congênitas pelo sistema do SUS

Dayane Gonçalves Hermam¹
José Luiz Padilha da Silva²

Resumo

O desenvolvimento de anomalias congênitas durante o período gestacional pode ocorrer por motivos diversos. O objetivo desse trabalho é estudar a base pública do Sistema Único de Saúde do Brasil (SUS) e observar as variáveis que podem ter alguma relação com o desenvolvimento dessas anomalias, aplicando um modelo linear generalizado para comprovação de hipótese na base em estudo.

Palavras-chave: Anomalia congênita, Modelo Linear Generalizado, SUS, SINASC, Distribuição binomial, Modelo logístico, R.

Abstract

The congenital anomalies are bad formations of any kind during the pregnancy, and can occur for many reasons. The object in this article is analyze the public base of the Brazilian System Unique of Health (SUS), and study the variables that could be related to these congenital anomalies developments, applying the statistical model called generalized linear model to validate the hypotheses raised.

Keywords: congenital anomalies, Generalized Linear Model, SUS, SINASC, Binomial distribution, Logistic model, R.

1 Introdução

A anomalia congênita é qualquer tipo de má formação que ocorre durante uma gestação e pode ter as mais diversas razões, desde a alimentação da gestante até o acesso que essa gestante tem a auxílios quando necessário, bem como acompanhamento médico durante toda a gestação. Dependendo da anomalia desenvolvida, essa criança pode ter uma dependência de cuidados temporários podendo se estender por todo o seu tempo de vida.

Existem hoje programas como o Planifica SUS que tem por objetivo regionalizar o atendimento do SUS, tentando assim identificar a necessidade de cada região e assim determinar as melhorias necessárias para cada necessidade. Para esse plano, o estado do Rio Grande

do Sul emitiu uma nota técnica[1] para tratar especificamente sobre o acompanhamento de gestantes e puérperas, onde visa além de promover a diminuição dos índices de mortes materna e infantil evitáveis, prever o acompanhamento de pacientes com doenças crônicas, que precisam de acompanhamento mais próximo, tentando garantir assim que os cuidados necessários serão tomados quando, da forma e no tempo necessário.

Nesse artigo, vamos estudar a seguinte questão: ao analisar os dados da mãe e da gestação em si, bem como histórico de gestações anteriores, podemos encontrar padrões que possam identificar o risco de desenvolvimento dessas anomalias previamente? A proposta desse trabalho é explorar a base do Sistema Único de Saúde (SUS) para tentar responder a essa pergunta, e levantar mais questionamentos na relação entre dados da mãe e de suas gestações com a possibilidade de desenvolvimento de anomalias congênitas.

2 Base de dados

Para realização deste estudo, vamos utilizar a base do SUS, como já foi citado. Para fins de volume de dados, vamos trabalhar com a base da seguinte forma:

- ▶ Temos 2.849.146 registros.
- ▶ Esses registros são referentes aos 26 estados do Brasil mais o Distrito Federal.
- ▶ O corte de tempo utilizado foi o período mais recente na base de dados, sendo, portanto, o ano de 2019.

A incidência de anomalias congênitas para o período citado é de 0,91%, os valores NA representam em torno de 6% para as variáveis a serem estudadas. A base do SUS apresenta muitas informações relacionadas aos pais, gestação e histórico gestacional da mãe, mas para esse estudo escolhemos apenas algumas para análise, pois consideramos que as demais poderiam ser colineares ou não agregar algum valor real ao estudo.

2.1 Pré-processamento

A única técnica de pré-processamento utilizada nesse trabalho, foi utilizando o pacote Microdatsus[2], que não é um pacote padrão da biblioteca do R. Esse pacote valida,

¹Aluno do programa de Especialização em Data Science & Big Data, day.gherman@hotmail.com.

²Professor do Departamento de Estatística - DEST/UFPR.

agrupa e categoriza algumas informações e insere algumas descrições que podem vir a facilitar o entendimento das informações ali presentes. Os nomes utilizados nesse trabalho, seguem as orientações de nomenclatura disponíveis no dicionário de dados do DataSUS[3]. Após o pré-processamento, escolhemos as nossas variáveis de interesse, que foram divididas em 4 grupos:

- ▶ **Por região:** Aqui escolhemos analisar por estado, pois por cidade a diferença não era muito evidente, e por região os indicadores poderiam perder a significância. Os estados com maiores indicadores poderiam perder relevância se analisados em conjunto com estados com menor indicador.
- ▶ **Informações dos pais:** Os indicadores escolhidos para esse grupo foram: *Idade da mãe*, *Idade do pai* e *Estado civil da mãe*; existem na base de dados outras variáveis referente a informações dos pais, mas foram desconsideradas para esse estudo, pois seriam redundantes com as informações já presentes.
- ▶ **Histórico gestacional da mãe:** Poderia o histórico gestacional da mãe indicar um possível desenvolvimento de anomalia? Para tentar responder a essa pergunta, as variáveis a serem analisadas são: *Quantidade de gestações anteriores*, *Quantidade de filhos nascidos vivos*, *Quantidade de filhos nascidos mortos/abortos*, *Quantidade de partos normais* e *Quantidade de partos cesárea*.
- ▶ **Informações da gestação atual e do bebê:** Aqui consideramos os dados que temos sobre a gestação atual precedendo ao momento do parto. As variáveis selecionadas nesse grupo são: *Quantidade de consultas pré-natal realizadas*, *Mês início do pré-natal*, *Tipo de gravidez* e *Sexo do bebê*.

3 Método estatístico

Vamos utilizar um mapa para realizar a primeira análise de incidência por estado e gráficos de barra para as demais variáveis. Vamos utilizar tabelas para entendermos a proporção de ocorrência de evento para cada variável analisada. O modelo a ser aplicado nesse estudo é o Modelo Linear Generalizado[4] com distribuição binomial (modelo logístico). Esse modelo foi escolhido pois, nossa variável resposta é uma variável binária - há desenvolvimento de anomalia, ou não há desenvolvimento de anomalia?

Para desenvolvimento desses estudos, utilizamos o R versão 4.0.4, os pacotes microdatasus (externo)[2], sf, geobr (externo), ggpubr, ggplot2, dplyr e tidyverse. O nível de significância foi de 5%.

4 Resultados e Discussões

4.1 Por região

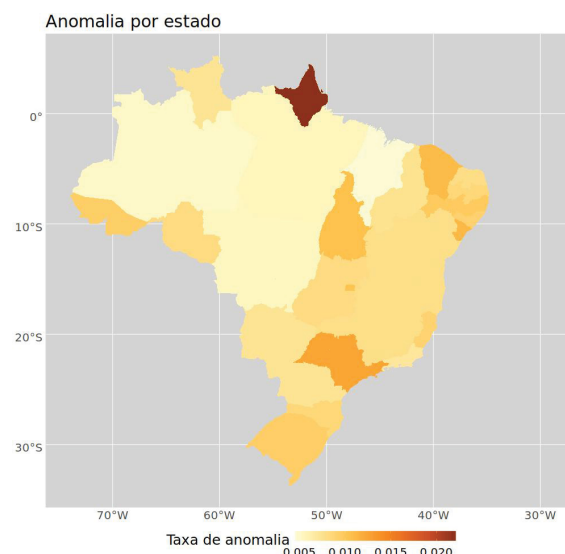


Figura 1: Mapa de incidência de desenvolvimento de anomalia congênita por estado.

	Total Reg	Ocorr	%		Total Reg	Ocorr	%
Acre	16.028	150	0,936%	Paraíba	57.329	489	0,853%
Alagoas	49.468	447	0,904%	Paraná	152.463	1.109	0,727%
Amapá	14.449	319	2,208%	Pernambuco	132.708	1.293	0,974%
Amazonas	77.473	398	0,514%	Piauí	47.582	356	0,748%
Bahia	193.287	1.510	0,781%	Rio de Janeiro	203.293	1.411	0,694%
Ceará	121.914	1.336	1,096%	Rio Grande do Norte	42.984	340	0,791%
Distrito Federal	35.372	365	1,032%	Rio Grande do Sul	132.339	1.244	0,940%
Espírito Santo	54.640	490	0,897%	Rondônia	26.293	214	0,814%
Goias	91.516	744	0,813%	Roraima	14.579	106	0,727%
Maranhão	112.075	510	0,455%	Santa Catarina	97.781	836	0,855%
Mato Grosso	58.647	318	0,542%	São Paulo	580.345	7.244	1,248%
Mato Grosso do Sul	43.375	316	0,729%	Sergipe	32.466	361	1,112%
Minas Gerais	246.564	1.922	0,780%	Tocantins	24.333	257	1,056%
Pará	136.138	753	0,553%	Total:	1.250.946	9.578	0,766%

Figura 2: Ocorrência por estado.

Como podemos ver na Figura 1, quando observamos a taxa de incidência por estado, alguns estados apresentam grande destaque. Observando o Estado do Amapá, por exemplo, podemos observar que o índice de incidência é bem alto, se compararmos com os demais estados da região norte.

Na Figura 2 podemos observar a quantidade de registros, ocorrências e percentual de incidência de desenvolvimento de anomalia congênita por estado. Essa tabela pode nos ajudar a ter melhor visão sobre os dados que compõe o mapa da Figura 1.

4.2 Informações dos pais

Nesse espaço, vamos olhar rapidamente para as variáveis que temos da mãe e do pai, informações que possam nos mostrar alguma tendência de desenvolvimento de doenças.

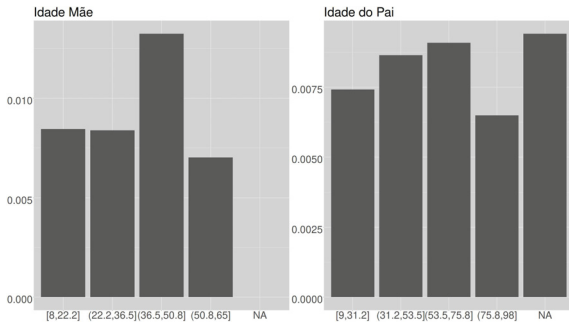


Figura 3: Gráficos de barras representando a média de incidência por faixas de idade. No primeiro gráfico, idade da mãe, no segundo gráfico, idade do pai.

Nesse caso podemos ver claramente nas Figuras 3 e 5 que na idade do pai há muitas informações como NA, mas não faz sentido ignorarmos essa informação, pois provavelmente um grande volume de gestantes dá a luz à crianças com pai não declarado. Mas podemos observar que a idade da mãe pode ter muito mais relação com o desenvolvimento de alguma anomalia que a idade do pai. Esse cenário pode mudar caso olhemos para os dados por estado.

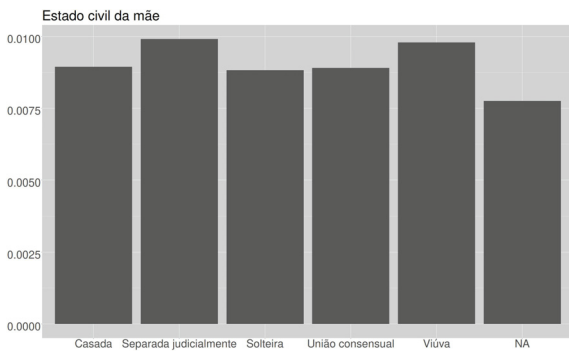


Figura 4: Gráfico de barras representando a média de incidência por estado civil da mãe.

Idade da mãe	Total Reg	Ocorr	%	Idade do pai	Total Reg	Ocorr	%
Menos de 18 anos	191.091	1.733	0,907%	Menos de 18 anos	8.976	75	0,836%
Entre 18 e 29 anos	1.565.876	12.676	0,810%	Entre 18 e 29 anos	423.357	3.105	0,733%
Entre 30 e 39 anos	943.539	8.805	0,933%	Entre 30 e 39 anos	475.924	3.854	0,810%
Entre 40 e 49 anos	94.461	1.620	1,715%	Entre 40 e 49 anos	133.350	1.285	0,964%
50 anos ou mais	407	4	0,983%	50 anos ou mais	21.567	211	0,978%
Total	2.795.374	24.838	0,889%	Total	1.063.174	8.530	0,802%

Estado Civil mãe	Total Reg	Ocorr	%
Casada	920.174	8.225	0,894%
Separada judicialmente	38.058	377	0,991%
Solteira	1.259.869	11.119	0,883%
União consensual	548.749	4.886	0,890%
Viúva	4.595	45	0,979%
Total	2.771.445	24.652	0,889%

Figura 5: Ocorrência por variáveis dos pais.

Na Figura 4 podemos observar a distribuição de incidência, considerando o estado civil da mãe. Por alguma razão, gestantes que estão nos grupos "Separada judicialmente" e "Viúva", apresentam um indicador um pouco mais alto que as demais para desenvolvimento de algum tipo de anomalia.

A Figura 5 nos ajuda a entender melhor os valores apresentador nas Figuras 3 e 4, mostrando a quantidade

de registros, ocorrências e percentual de eventos para cada variável citada anteriormente nessa seção.

4.3 Histórico gestacional da mãe

Nesse tópico, vamos observar as mães que já tiveram outra(s) gestação(ões), como se deu e qual a incidência de desenvolvimento de anomalia por essas mães nas gestações correntes.

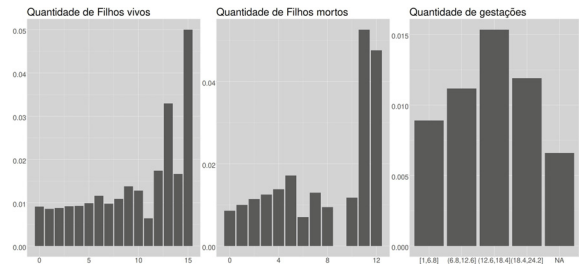


Figura 6: Gráficos mostrando distribuição de taxa de anomalia desenvolvida comparado com informações de gestações anteriores. No primeiro gráfico dados de quantidade de filhos nascidos vivos, segundo gráfico representando quantidade de abortos e filhos nascidos mortos e no terceiro gráfico a quantidade de gestações anteriores.

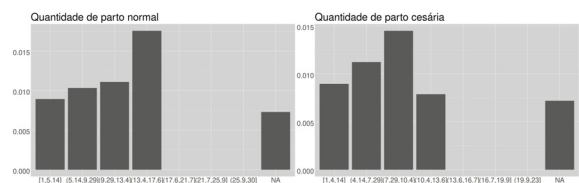


Figura 7: Gráficos demonstrando a distribuição de ocorrências por tipo de parto; primeiro gráfico demonstrando partos normais, segundo gráfico partos cesária

Qtd Filhos mortos	Total Reg	Ocorr	%	Qtd Filhos Vivos	Total Reg	Ocorr	%
0	2.165.934	18.596	0,859%	0	1.126.301	10.346	0,919%
1	421.646	4.218	1,000%	1	902.136	7.802	0,865%
2	79.794	910	1,140%	2	405.590	3.590	0,885%
3	17.564	220	1,253%	3	152.390	1.410	0,925%
4 ou mais	6.893	95	1,378%	4 ou mais	120.887	1.197	0,990%
Total	2.691.831	24.039	0,893%	Total	2.707.304	24.345	0,899%

Qtd Gest Ant	Total Reg	Ocorr	%	Qtd Parto Ces	Total Reg	Ocorr	%
1	996.761	8.837	0,89%	1	1.898.697	16.881	0,89%
2	847.699	7.214	0,85%	2	592.321	5.294	0,89%
3	466.526	4.215	0,90%	3	156.460	1.483	0,95%
4 ou mais	408.709	4.072	1,00%	4 ou mais	35.518	374	1,04%
Total	2.719.695	24.338	0,89%	Total	2.683.396	24.032	0,90%

Qtd Parto Norm	Total Reg	Ocorr	%
1	1.737.887	15.698	0,90%
2	513.384	4.459	0,87%
3	235.822	2.021	0,86%
4 ou mais	208.972	1.935	0,93%
Total	2.696.065	24.113	0,89%

Figura 8: Ocorrência por variáveis de histórico gestacional anterior.

Quando vemos por tipo de parto (Figuras 7 e 8), podemos observar que o gráfico apresenta uma tendência crescente nas incidências diretamente relacionada com a quantidade de partos de cada modalidade descrita, mas

nesse caso fica difícil dizermos se a relação é com o tipo de parto anterior ou apenas com a quantidade de partos realizados previamente. Mas vemos claramente, pelo menos no primeiro gráfico, o qual representa partos normais, que quanto mais processos de parto a mãe passou aparentemente maior as chances de desenvolvimento de algum tipo de anomalia. Para o gráfico de partos cesária, vemos uma queda na ultima barra antes das informações de NA.

Na Figura 8 podemos ver a distribuição de registros e ocorrência e o percentual para cada variável apresentava neste grupo de variáveis.

4.4 Informações da gestação atual e do bebê

Nesta seção vamos observar as variáveis relacionadas com a gestação atual e do bebê. Nesse caso, consideramos apenas informações de períodos anteriores e durante a gestação atual; informações parto e pós-parto foram desconsideradas.

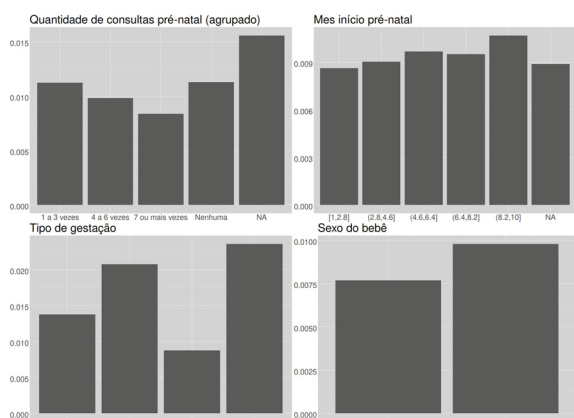


Figura 9: Combinação com 4 gráficos de informações sobre o acompanhamento da gestação e do bebê; primeiro gráfico demonstra distribuição do indicador por quantidade de consultas pré-natal realizadas durante a gestação; segundo gráfico apresenta o mês de início do acompanhamento do pré-natal; terceiro gráfico representa por tipo de gravidez e o último gráfico demonstra o desenvolvimento de anomalias por sexo do bebê

Tipo de gravidez	Total Reg	Ocorr	%
Dupla	59.527	819	1,376%
Tripla e mais	1.445	30	2,076%
Única	2.732.730	23.948	0,876%
Total	2.793.702	24.797	0,888%

Qtd Cons pré-natal	Total Reg	Ocorr	%
1 a 3 vezes	148.906	1.672	1,123%
4 a 6 vezes	565.497	5.554	0,982%
7 ou mais vezes	2.027.774	16.962	0,836%
Nenhuma	41.762	471	1,128%
Total	2.783.939	24.659	0,886%

Mês início pré-natal	Total Reg	Ocorr	%
Primeiro mês	608.438	5.044	0,829%
Segundo mês	1.062.702	9.378	0,882%
Terceiro mês	525.978	4.650	0,884%
Quarto mês	226.672	2.156	0,951%
Quinto mês	120.242	1.153	0,959%
Sexto mês	61.943	612	0,988%
Sétimo mês	32.556	318	0,977%
Oitavo mês ou mais	88.767	921	1,038%
Total	2.727.298	24.232	0,888%

Sexo do bebê	Total Reg	Ocorr	%
Feminino	1.365.166	10.492	0,769%
Masculino	1.429.925	13.998	0,979%
Total	2.795.091	24.490	0,876%

Figura 10: Ocorrência por variáveis de gestação atual e bebê.

No primeiro gráfico da Figura 9 podemos perceber a possível relação direta e inversa entre quantidade de

consultas pré-natal realizadas pela mãe com o desenvolvimento de algum tipo de anomalia. Já no segundo gráfico da Figura 9 vemos uma informação complementar a informação anterior que fala sobre o mês que a mãe iniciou os cuidados com pré-natal com o desenvolvimento de anomalia. Podemos ver que mães que iniciam o acompanhamento mais tarde apresentam uma chance de desenvolvimento de anomalia aumentada em relação aos demais grupos. No terceiro gráfico, observamos que gestações múltiplas podem ter mais chances de desenvolver algum tipo de anomalia, e por fim, vemos uma distribuição por sexo do bebê. A diferença não é muito expressiva nesse caso, mas vamos considerar a variável no modelo. Para o último gráfico da Figura 9, vale ressaltar que desconsideramos os valores NA, pois com eles era impossível observar as barras por sexo.

A Figura 10 representa a distribuição de registros e ocorrência de desenvolvimento de anomalia congênita por variável deste grupo.

4.5 Modelo Simples

Após ajustarmos o modelo simples, identificamos que algumas variáveis não se mostram muito relacionadas com a variável resposta e foram desconsideradas nas etapas seguintes. Na Figura 11, podemos observar o resultado dessa fase:

	LRT	Pr(>Chi)
Estado	2.040,1831	0,0000
Idade da mãe	372,2119	0,0000
Idade do pai	67,7384	0,0000
Estado civil da mãe	5,5563	0,2348
Qtd de Gestações anteriores	54,6030	0,0000
Qtd Filhos nascidos vivos	4,7330	0,0296
Qtd de filhos mortos e/ou abortos	156,0323	0,0000
Qtd Parto normal	0,4455	0,5045
Qtd parto cesária	9,5448	0,0020
Qtd consultas pré-natal	228,4869	0,0000
Mês início pré-natal	69,1672	0,0000
Tipo de gravidez	158,8957	0,0000
Sexo	357,4284	0,0000

Figura 11: Estatística da razão de verossimilhanças e valor-p para cada variável no modelo simples.

Com base nos resultados acima, foram desconsideradas as seguintes variáveis para o prosseguimento das análises: *Estado civil da mãe, Quantidade de filhos vivos, Quantidade de parto normal, Quantidade de parto cesária.*

4.6 Modelo múltiplo

Para o modelo final são consideradas apenas as variáveis estatisticamente significativas ao nível de 5%. As Figuras 12 e 13 demonstram o resultado para cada variável:

		Modelo Simples				Modelo Múltiplo			
		Odds ratio	LI	LS	p value	Odds ratio	LI	LS	p value
Estado	Acre	Ref	-	-	-	Ref	-	-	-
	Alagoas	0,965	0,802	1,162	0,709	1,386	1,095	1,755	0,007
	Amapá	2,390	1,966	2,905	0,000	3,046	2,390	3,883	0,000
	Amazonas	0,547	0,453	0,660	0,000	0,693	0,546	0,878	0,002
	Bahia	0,833	0,704	0,986	0,034	1,114	0,893	1,389	0,337
	Ceará	1,173	0,990	1,390	0,065	1,551	1,242	1,938	0,000
	Distrito Federal	1,104	0,912	1,336	0,311	1,435	1,131	1,822	0,003
	Espírito Santo	0,958	0,797	1,151	0,646	1,308	1,037	1,649	0,023
	Goiás	0,868	0,727	1,035	0,114	1,154	0,920	1,447	0,216
	Maranhão	0,484	0,403	0,581	0,000	0,662	0,525	0,835	0,001
	Mato Grosso	0,577	0,475	0,701	0,000	0,792	0,622	1,008	0,058
	Mato Grosso do Sul	0,777	0,639	0,944	0,011	1,054	0,828	1,342	0,667
	Minas Gerais	0,832	0,704	0,983	0,030	1,144	0,920	1,424	0,227
	Pará	0,589	0,494	0,702	0,000	0,783	0,624	0,983	0,035
	Paraíba	0,911	0,758	1,094	0,318	1,279	1,014	1,613	0,038
	Paraná	0,776	0,654	0,921	0,004	1,097	0,878	1,369	0,415
	Pernambuco	1,041	0,879	1,234	0,639	1,418	1,137	1,768	0,002
	Piauí	0,798	0,659	0,966	0,021	0,956	0,749	1,220	0,718
	Rio de Janeiro	0,740	0,625	0,876	0,000	1,003	0,805	1,250	0,978
	Rio Grande do Norte	0,844	0,696	1,024	0,085	1,161	0,911	1,479	0,228
	Rio Grande do Sul	1,004	0,847	1,191	0,959	1,353	1,085	1,688	0,007
	Rorônia	0,869	0,704	1,071	0,188	0,933	0,711	1,223	0,614
	Roraima	0,775	0,604	0,995	0,046	0,919	0,681	1,241	0,583
	Santa Catarina	0,913	0,767	1,087	0,306	1,268	1,013	1,587	0,038
	São Paulo	1,338	1,137	1,574	0,000	1,825	1,472	2,263	0,000
	Sergipe	1,190	0,983	1,441	0,074	1,626	1,277	2,071	0,000
	Tocantins	1,130	0,923	1,383	0,237	1,568	1,224	2,007	0,000

Figura 12: Resultado final por estado.

		Modelo Simples				Modelo Múltiplo			
		Odds ratio	LI	LS	p value	Odds ratio	LI	LS	p value
Idade da mãe	Idade	1,018	1,016	1,020	0,000	1,018	1,016	1,020	0,000
	Quantidade	1,131	1,111	1,152	0,000	1,142	1,117	1,167	0,000
Qtd de filhos mortos e/ou abortos	Nenhuma	Ref	-	-	-	Ref	-	-	-
	1 a 3 vezes	0,996	0,898	1,103	0,932	1,162	0,997	1,355	0,055
	4 a 6 vezes	0,870	0,791	0,956	0,004	0,995	0,856	1,156	0,948
	7 ou mais vezes	0,740	0,674	0,811	0,000	0,782	0,672	0,909	0,001
Mês início pré-natal	Mês	1,029	1,022	1,036	0,000	1,027	1,018	1,035	0,000
	Única	Ref	-	-	-	Ref	-	-	-
Tipo de Gravidez	Dupla	1,578	1,471	1,693	0,000	1,474	1,370	1,585	0,000
	Tripla e mais	2,398	1,670	3,444	0,000	2,167	1,489	3,133	0,000
	Sexo	Masculino	1,276	1,244	1,309	0,000	1,282	1,249	1,316

Figura 13: Resultado final para as demais variáveis

Note que na Figura 13, a variável de *Idade do pai* está ausente, isso se deve pois quando aplicamos o modelo múltiplo, ela se mostrou não significativa.

Na Figura 12, podemos observar o resultado por estado, considerando o Acre como valor de referência; sendo assim, a interpretação dos resultados fica da seguinte forma observando a coluna de valores Modelo Simples: Uma gestante do estado do Alagoas tem chances diminuídas em 0,035% se comparada à uma gestante do estado do Acre; já uma gestante do Amapá, tem as chances aumentadas em 139%, se comparado ao mesmo estado de referência. Na Figura 13 podemos observar algumas outras variáveis que se leem da mesma forma, sendo elas: *Quantidade de consultas pré-natal*, *tipo de gravidez* e *sexo do bebê*.

Ainda na Figura 13, podemos interpretar os resultados das demais variáveis da seguinte forma: Para Idade da mãe, por exemplo, temos a seguinte leitura: Para cada aumento de um ano de idade da mãe, as chances de ocorrência do evento são aumentadas em 1,8%, com intervalo de confiança entre 1,6% e 2,0%, a mesma lógica se aplica para as seguintes variáveis: *Quantidade de filhos nascidos mortos e abortos*, *Quantidade de gestações anteriores* e *Mês início do pré-natal*.

Note que na Figura 13 os valores apresentam pouca variação de valores de *Odds* entre os modelos, isso quer dizer que o efeito do Estado não altera as conclusões para as variáveis de mãe e parto. Já na Figura 12 podemos observar que alguns efeitos mudam a direção e significância da variável. Isso significa que as variáveis

são diferentes para os estados. Por exemplo, a idade da mãe em Alagoas e São Paulo não devem ser semelhantes.

Das 13 variáveis apresentadas inicialmente, 8 chegaram à etapa final e pudemos confirmar algumas hipóteses levantadas no início desse trabalho, porém, o poder discriminatório e o intervalo de confiança são muito baixos como podemos observar na Figura 14:

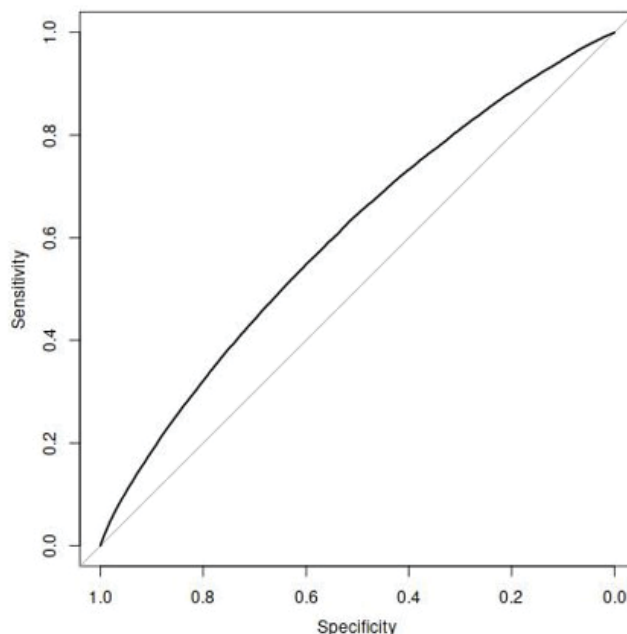


Figura 14: Área sob a curva ROC

Para melhorar o poder de discriminação do modelo, algumas abordagens diferentes podem ser trabalhadas, como o aumento de variáveis analisadas, podendo ser desde variáveis presentes na base do SUS como variáveis externas, como dados sobre renda que temos no IBGE, ou até mesmo a base de dados do INSS. Pode-se ainda considerar um período de tempo mais amplo, pois o evento estudado tem uma incidência muito baixa na base de dados, sendo assim, o percentual de ocorrências deve se manter, mas a quantidade de eventos a serem analisados deve aumentar. E por fim, pode-se também aplicar outros modelos na base de dados e comparar o desempenho de cada modelo. Seria interessante também a realização de trabalhos futuros em conjunto com profissionais da área de obstetria.

5 Conclusões

Existem hoje poucos estudos recentes acerca do assunto do desenvolvimento de anomalias congênitas, e acredito que esse trabalho se encerra com sucesso ao que se propôs, o de estudar a base de dados e identificar possível relação entre as informações presente na base de dados do SUS e o desenvolvimento de anomalias congênitas, aplicando os conhecimentos adquiridos no curso

da especialização. Espero que esse estudo seja de alguma valia para a comunidade, mesmo sendo apenas como objeto de estudo das técnicas aplicadas ou até mesmo dando continuidade a esse estudo. Acredito que existe muita informação que pode ser descoberta através da análise de dados públicos, e assim gerar algum impacto positivo, por menor que seja, na vida da comunidade.

Agradecimentos

Gostaria de agradecer primeiramente ao meu orientador, Prof. Dr. José Luiz Padilha da Silva, por toda a paciência e tempo dedicado ao desenvolvimento desse trabalho, sem sua ajuda, esse trabalho jamais teria evoluído como evoluiu. Obrigada por todo o apoio. Agradecimento especial a toda a equipe da UFPR com quem tive contato, professores do departamento de estatística e do departamento de informática que guiaram a turma pelo último ano e meio, com desafio extra lecionando remotamente para que não perdessemos o curso. Também gostaria de agradecer à equipe de suporte do Departamento de Informática da UFPR, pois sem o auxílio deles para configurar o servidor para acesso remoto, eu não teria conseguido processar todo esse volume de dados.

Referências

[1] SAÚDE DA MULHERNA GESTAÇÃO, PARTO E PUERPÉRIO (*Panifica SUS RS*)
<https://atencao basica.saude.rs.gov.br/upload/arquivos/202001/03091259-nt-gestante-planificasus.pdf>

[2] SALDANHA, Raphael de Freitas; BASTOS, Ronaldo Rocha; BARCELLOS, Christovam. *Microdatasus: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS)*. *Cad. Saúde Pública, Rio de Janeiro*, v. 35, n. 9, e00032419, 2019. Available from <http://ref.scielo.org/dhcq3y>.

[3] *Dicionário de dados SINASC*
<http://vigilancia.saude.mg.gov.br/index.php/download/dicionario-de-dados-da-tabela-dn-dasisms/?wpdmdl=919>
Estudo de prevalência de defeitos congênitos no Vale do Paraíba Paulista (2000 - 2003):
<https://www.scielo.br/j/rpp/a/NnQJrN4jSFxwMpHqRP7Fscq/?lang=pt>

[4] *Modelos Lineares Generalizados*
<http://leg.ufpr.br/lucambio/GLM/GLM.html>