

**Universidade de São Paulo
Instituto de Matemática e Estatística**

Centro de Estatística Aplicada

Relatório de Análise Estatística

RAE-CEA-21P17

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

“Associação entre medidas eletrofisiológicas e de neuroimagem funcional com a resposta ao tratamento farmacológico de primeira linha em pacientes com transtorno obsessivo-compulsivo ”

Carlos Alberto de Braganca Pereira

Ivan Colagrande Castro

Vitor Hugo Vieira de Lima

São Paulo, dezembro de 2021

CENTRO DE ESTATÍSTICA APLICADA - CEA – USP

TÍTULO: Associação entre medidas eletrofisiológicas e de neuroimagem funcional com a resposta ao tratamento farmacológico de primeira linha em pacientes com transtorno obsessivo-compulsivo

PESQUISADORES: Daniel Lucas Conceição Costa e Juliana Belo Diniz

INSTITUIÇÃO: Instituto de Psiquiatria do Hospital das Clínicas da FMUSP

FINALIDADE DO PROJETO: Publicação

RESPONSÁVEIS PELA ANÁLISE: Carlos Alberto de Braganca Pereira

Ivan Colagrande Castro

Vitor Hugo Vieira de Lima

REFERÊNCIA DESTE TRABALHO: PEREIRA, C. A. B.; CASTRO, I.C.; LIMA, V. H. V. **Relatório de análise estatística sobre o projeto: “Associação entre medidas eletrofisiológicas e de neuroimagem funcional com a resposta ao tratamento farmacológico de primeira linha em pacientes com transtorno obsessivo-compulsivo”**. São Paulo, IME-USP, 2021. (RAE–CEA-21P17)

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

BUSSAB, W.O.; MORETTIN, P.A (2017). **Estatística Básica**. 9. ed. São Paulo: Editora Saraiva

COOK, R.D.; WEISBERG, S. (1982). **Residuals and Influence in Regression**. 1.ed. New York: Chapman and Hall.

GOODMAN, W.K.; PRICE, L.H.; RASMUSSEN, S.A. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. **Arch Gen Psychiatry**, **46**.

HOLLANDER, E. (1997). Obsessive-compulsive disorder: the hidden epidemic. **J Clin Psychiatry**. **58**: 3-6.

HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R.X. (2013). **Applied Logistic Regression**. 3. ed. New York: John Wiley. 177 p.

KESSLER, R.C.; BERGLUND, P.; DEMLER, O.; JIN, R.; WALTERS, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. **Arch Gen Psychiatry**, **62**,593-602.

LUND, B.; BROTHERTON, D. (2013) Information Value Statistic. **Division of Marketing Associates**, Detroit, MI.

PERICCHI, L.; PEREIRA, C. (2016). Adaptative significance levels using optimal decision rules: Balancing by weighting the error probabilities. **Brazilian Journal of Probability and Statistics**, **30**, 70–90.

RUSCIO, A.M.; STEIN, D.J.; CHIU, W.T.; KESSLER, R.C. (2010). **The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication**. *Mol Psychiatry*. 53-63.

THOMAS, L.C.; EDELMAN, D.B.; CROOK, J.N. (2002). **Credit Scoring and its Applications**. 1, ed. , Philadelphia, PA

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Microsoft Word for Windows (versão 2016)

Microsoft Excel for Windows (versão 2016)

R for Windows, versão R-4.0.3

Rstudio for Windows, versão 1.3.1093

Power BI Desktop, 2.96.701.0 64-bit (agosto de 2021)

TÉCNICAS ESTATÍSTICAS UTILIZADAS:

03:010 – Análise Descritiva Unidimensional

03:020 – Análise Descritiva Multidimensional

03:990 – Outros

05:010 – Testes de Hipóteses Paramétricas

05:070 – Testes de Hipóteses Não Paramétricas

08:010 – Análise de Variância com Efeitos Fixos

ÁREA DE APLICAÇÃO:

Psicometria (14:090)

Resumo

O transtorno obsessivo-compulsivo (TOC) é um transtorno mental caracterizado pela presença de compulsões e obsessões, que podem ser debilitantes para os portadores. Foi realizado um experimento de medo condicionado em 37 pacientes, e aferidas medidas de condutância de pele através de choque. O objetivo é de encontrar possíveis marcadores e preditores da melhora dos pacientes (critério previamente definido) e da redução da escala YBOCS, dentro das variáveis que nos foram providas.

Dadas as duas possibilidades de variável resposta descritas acima, foram ajustados três tipos de modelo: de regressão linear para a redução de YBOCS, e análise discriminante e regressão logística para a melhora dos pacientes. Os candidatos a preditores e marcadores foram selecionados via *Informative Value*, sendo complementados pelas variáveis requeridas pelos pesquisadores.

Para preditores de obsessão, destacam-se a presença de TOC de contaminação, a gravidade do TOC sexual e a gravidade de ansiedade. Para preditores de compulsão, destacam-se a gravidade do TOC sexual, o sexo e a qualidade de vida geral do paciente. Para marcadores de obsessão, destacam-se a diferença entre os estímulos com e sem choque não extintos, e as variáveis pós-choque 1 e pós-choque 9. Para marcadores de compulsão, destacam-se pós-choque 1, pós-choque 9 e pós-choque 3.

Sumário

1. Introdução	8
2. Objetivo do estudo	9
3. Descrição de estudo	9
4. Descrição das variáveis	11
5. Análise descritiva	14
5.1. Variável resposta	14
5.2. Seleção de variáveis	17
5.3. Análise descritiva das variáveis selecionadas	20
6. Análise Inferencial	31
6.1. Variável Numérica	31
6.1.1. Regressão linear para obsessão	32
6.1.2. Regressão linear para compulsão	41
6.2. Variável Categórica	46
6.2.1. Regressão logística para Obsessão	46
Preditores	Erro! Indicador não definido.
Marcadores	Erro! Indicador não definido.
6.2.2. Regressão logística para Compulsão	51
6.2.3. Análise Discriminante para obsessão e compulsão	Erro! Indicador não definido.
7. Considerações Finais	65

1. Introdução

O transtorno obsessivo-compulsivo (TOC), é um transtorno mental grave, muitas vezes incapacitante e que pode provocar prejuízo em diferentes esferas da vida dos portadores, como trabalho, estudo e relações interpessoais. Caracteriza-se pela presença de obsessões e/ou compulsões. Obsessões são pensamentos, ideias ou imagens intrusivos, repetitivos, com conteúdo desagradável e que, por isso, causam incômodo. Já as compulsões são comportamentos repetitivos ou rituais mentais realizados de forma estereotipada com o intuito de aliviar o desconforto causado pelas obsessões. Como as compulsões provocam um alívio temporário desse desconforto, estes comportamentos tendem a se repetir. Os principais temas dos sintomas relacionam-se à violência/agressão, sujeira/contaminação, sexo ou religião, simetria/organização e acumulação.

A prevalência do TOC ao longo da vida situa-se entre 2-3% da população geral, enquanto a sua incidência em 12 meses é 1,2% (Kessler, 2005; Ruscio, 2008). É importante ressaltar, inclusive para os objetivos deste estudo, que o TOC acomete de modo desigual pessoas de faixas etárias diferentes - na infância, a prevalência do TOC é maior em meninos do que em meninas. Em adultos, a prevalência do TOC entre os sexos é semelhante.

Frequentemente, pacientes com TOC apresentam diagnósticos psiquiátricos adicionais, as chamadas comorbidades - depressão, transtornos de ansiedade, transtorno de déficit de atenção e transtornos do espectro do TOC são diagnosticados em 30-90% dos portadores de TOC. Além disso, os sintomas costumam aparecer de forma insidiosa e pioram ao longo do tempo, o que pode dificultar o diagnóstico precoce. Estima-se que os pacientes podem levar até 17 anos para receber tratamento específico para o TOC (Hollander, 1997). Fatores que contribuem para este atraso incluem o desconhecimento sobre o transtorno e a vergonha que os portadores sentem por causa dos seus sintomas, já que na maioria dos casos, os pacientes reconhecem o exagero das preocupações obsessivas e dos rituais, mas não conseguem evitá-los.

Apesar de existirem tratamentos de primeira linha para o TOC, incluindo os medicamentos inibidores seletivos da recaptação de serotonina (ISRS) e a terapia cognitivo-comportamental (TCC), é comum a ausência de resposta terapêutica.

Estima-se que aproximadamente metade dos pacientes com TOC tratados com uma dessas opções de primeira linha não apresente melhora clinicamente significativa.

Existe uma extensa literatura a respeito de características clínicas e sócio-demográficas consideradas como preditoras de resposta ao tratamento farmacológico inicial do TOC, mas a identificação de características de neuroimagem ou medidas eletrofisiológicas que possam prever de forma confiável a resposta ao tratamento com os ISRS constitui uma lacuna na literatura.

2. Objetivo do estudo

Tendo em vista o cenário explicitado na seção anterior, o presente estudo se propõe a analisar os dados de um experimento em que 37 pacientes com TOC e sem tratamento prévio, que foram submetidos a um experimento de medo condicionado e a um exame de ressonância magnética, antes e depois de 4 semanas de tratamento com um medicamento da família dos ISRS, que foi mantido em dose inalterada por em média 8 semanas adicionais.

A gravidade do TOC antes e ao longo do tratamento foi medida com a escala Yale-Brown Obsessive-Compulsive Scale (YBOCS) (Goodman et al, 1989). Esta escala é composta por 10 questões, das quais 5 se referem às obsessões e 5 às compulsões. As perguntas se referem ao tempo gasto, interferência na rotina, incômodo provocado, níveis de resistência e controle em relação às obsessões e compulsões. Cada pergunta tem 5 opções de resposta, que variam de 0 a 4, com escores menores indicando menor gravidade (e vice-versa).

O principal objetivo do estudo é identificar preditores e marcadores associados à resposta ao tratamento farmacológico de primeira linha para o TOC.

3. Descrição de estudo

Conforme relatado no texto principal dos pesquisadores, o recrutamento dos pacientes foi realizado no Instituto de Psiquiatria do Hospital das Clínicas da FMUSP. Os interessados passaram inicialmente por uma triagem com um psicólogo ou psiquiatra treinado, que realizava a seleção dos pacientes elegíveis ao estudo.

Os pacientes elegíveis deveriam ter idade entre 18 e 65 anos e uma medida de YBOCS superior a 16. Além disso, foram excluídos pacientes com distúrbios do

neurodesenvolvimento (autismo e deficiência intelectual), transtornos psicóticos, dependência química ou transtorno bipolar, que haviam feito uso de medicamentos psicoterápicos nas 12 semanas anteriores ao início do estudo, se recusaram a receber tratamento farmacológico, grávidas ou com risco de gravidez.

Após aplicados os critérios de seleção, foram incluídos 37 pacientes com TOC, que foram submetidos a um experimento de medo condicionado antes e depois de 4 semanas de tratamento com um medicamento inibidor seletivo da recaptação de serotonina (ISRS). O experimento consistia na apresentação de sequências de imagens com estímulos visuais diferentes (luz azul, amarela ou vermelha) e era realizado em 2 dias consecutivos. As diferentes fases do experimento se dividiam da seguinte forma: o condicionamento e a extinção do medo no primeiro dia e a recuperação e renovação no segundo.

Durante o condicionamento, os sujeitos eram condicionados a sentir medo de dois estímulos visuais diferentes, através do pareamento de choques elétricos de baixa intensidade nos dedos da mão direita aos estímulos visuais; na extinção, a resposta de medo ao estímulo condicionado era extinta por meio da apresentação de uma imagem diferente da usada no condicionamento e sem choque - nesta etapa, eram exibidos apenas dois estímulos, o que foi apresentado inicialmente sem choque e o que deveria ser extinto. Durante a recuperação, mantinha-se a imagem usada na extinção, acrescida do estímulo que permaneceu associado ao medo, porém sem choque. Por fim, a renovação apresentava o mesmo cenário utilizado no condicionamento, novamente sem a presença do choque.

Além do experimento de medo condicionado, que visava avaliar a persistência ou não da resposta do medo nos pacientes, medida através da variação da condutância de pele ao longo do experimento, foram coletados dados demográficos e clínicos, por meio de uma entrevista clínica estruturada e de neuroimagem, por meio de um exame de ressonância magnética.

4. Descrição das variáveis

- **Sexo:** Masculino e Feminino
- **Idade (anos)**
- **Estado civil:** Solteiro, Casado ou Amasiado, Divorciado ou Separado, Viúvo
- **Etnia:** Branca, Afrodescendente ou Oriental
- **Religião:** Católico, Evangélico, Protestante, Judeu, Espírita, Testemunha de jeová, Sem religião, Outras
- **Ocupação:** Trabalhando, Desempregado, Aposentado por tempo de serviço, Aposentado por invalidez, Afastado/licença médica, Do lar, Nunca trabalhou, Estudante
- **Escolaridade:** Analfabeto, Fundamental incompleto, Fundamental completo, Médio incompleto, Médio completo, Universitário incompleto, Universitário completo, Pós-graduação.
- **ABIPEME (Classe Social:)** A, B, C, D, E
- **História familiar de transtorno obsessivo-compulsivo:** Ausente, Primeiro grau, Segundo grau, Distante
- **História familiar de transtornos do espectro obsessivo-compulsivo:** Ausente, Primeiro grau, Segundo grau, Distante
- **História familiar de outros transtornos psiquiátricos:** Ausente, Primeiro grau, Segundo grau, Distante
- **Tratamento Passado:** Sem tratamento farmacológico anterior ou "drug-naive", Algum tratamento farmacológico anterior
- **Exposição fármacos:** Número de tratamentos farmacológicos anteriores
- **Refratariedade:** Número de tratamentos adequados anteriores completos e sem resposta (quanto maior, maior a refratariedade)

- **YBOCS obsessão:** Escala de gravidade para total de obsessões, medida em diversas semanas do estudo
- **YBOCS compulsões:** Escala de gravidade para total de compulsões, medida em diversas semanas do estudo
- **Comorbidade com transtorno de Humor:** Ausente, Presente
- **Gravidade de depressão (escore Inventário de Beck):** 0 a 63 (quanto maior, pior)
- **Comorbidade com Transtorno de Ansiedade:** Ausente, Presente
- **Gravidade de ansiedade (escore Inventário de Beck):** 0 a 63 (quanto maior, pior)
- **Qualidade de vida para saúde física:** 0 a 100
- **Qualidade de vida para limitação física:** 0 a 100
- **Qualidade de vida para problemas emocionais:** 0 a 100
- **Qualidade de vida para energia:** 0 a 100
- **Qualidade de vida para bem-estar:** 0 a 100
- **Qualidade de vida para social:** 0 a 100
- **Qualidade de vida para dor:** 0 a 100
- **Qualidade de vida para geral:** 0 a 100
- **Presença de sintomas de TOC do tipo agressão:** Ausente, Presente
- **Gravidade dos sintomas de TOC do tipo agressão:** 0 a 15 (quanto maior, pior)
- **Presença de sintomas de TOC do tipo sexual/religião:** Ausente, Presente
- **Gravidade dos sintomas de TOC do tipo sexual/religião:** 0 a 15 (quanto maior, pior)
- **Presença de sintomas de TOC do tipo simetria:** Ausente, Presente

- **Gravidade dos sintomas de TOC do tipo simetria:** 0 a 15 (quanto maior, pior)
- **Presença de sintomas de TOC do tipo contaminação:** Ausente, Presente
- **Gravidade dos sintomas de TOC do tipo contaminação:** 0 a 15 (quanto maior, pior)
- **Presença de sintomas de TOC do tipo acumulação:** Ausente, Presente
- **Gravidade dos sintomas de TOC do tipo acumulação:** 0 a 15 (quanto maior, pior)
- **Presença de sintomas de TOC do tipo diversas:** Ausente, Presente
- **Gravidade dos sintomas de TOC do tipo diversas:** 0 a 15 (quanto maior, pior)
- **Crítica sobre a doença, escore total da escala BABS:** 0 a 24 (quanto maior, pior)
- **Gravidade dos fenômenos sensoriais associados aos sintomas obsessivo compulsivos:** 0 a 15 (quanto maior, pior)
- **Choque:** Resultado da resposta de condutância da pele para o choque da semana 0, foram aplicados 10 choques em cada paciente, dessa forma temos uma medida para cada choque.
- **Pós Choque:** Resultado da resposta de condutância da pele para o choque na semana 4, também , foram aplicados 10 choques em cada paciente.

Para as próximas variáveis, temos a seguinte codificação:

- CS+E: Estímulo com choque que passará pelo processo de extinção
- CS-E: Estímulo extinto sem choque
- CS+U: Estímulo com choque que não passará pelo processo de extinção
- CS-U: Estímulo não extinto sem choque

Utilizando-se essa codificação, tem-se as variáveis:

- **Diferença de CS_E na recuperação:** Média da diferença entre a resposta ao CS+E menos a resposta ao CS-E na recuperação
- **Diferença de CS_U na recuperação:** Média da diferença entre a resposta ao CS+U menos a resposta ao CS-U na recuperação
- **Diferença de CS_E na renovação:** Média da diferença entre a resposta ao CS+E menos a resposta ao CS-E na renovação
- **Diferença de CS_U na renovação:** Média da diferença entre a resposta ao CS+U menos a resposta ao CS-U na renovação

5. Análise descritiva

Nesta seção será discutida a transformação da variável resposta selecionada para o projeto e o método adotado para seleção de candidatos a preditor e marcador com suas respectivas análises descritivas.

5.1. Variável resposta

A escala YBOCS é utilizada para quantificar a gravidade dos sintomas de compulsão e obsessão, seu valor foi medido em todas as semanas pré estabelecidas que o paciente retornou, porém como é possível observar na Tabela 1 a quantidade de respostas cai drasticamente após a semana 4, o que levou o estudo a se concentrar apenas nas semanas 0 e 4.

A Figura 1 apresenta os boxplots dos YBOCS obtidos na semana 0 e 4. Além da redução expressiva entre as semanas, se nota uma mudança na ordem das medianas de cada sintoma.

Para avaliar essa mudança, foram realizados testes de Kruskal Wallis e uma ANOVA de duas vias para as variáveis semana e sintoma contra a variável resposta. Em todos os casos não se obteve significância estatística para a variável sintoma (valor $p > 0.2$), apesar da diferença entre as medianas, principalmente na semana 0.

Após discussão com os pesquisadores, ficou acordado que, mesmo sem significância estatística, a análise deveria ser feita separadamente por sintoma, visto que a utilização de uma medida resumo para os dados de compulsão e obsessão poderia levar a tratar pacientes com gravidade distintas dos sintomas de forma igual.

Tabela 1: Quantidade de medições do YBOCS por semana do estudo

Semana	Quantidade
0	37
4	34
10	1
11	3
12	12
13	3
14	2
19	1
20	5
21	2
22	2
23	3
24	3
25	1

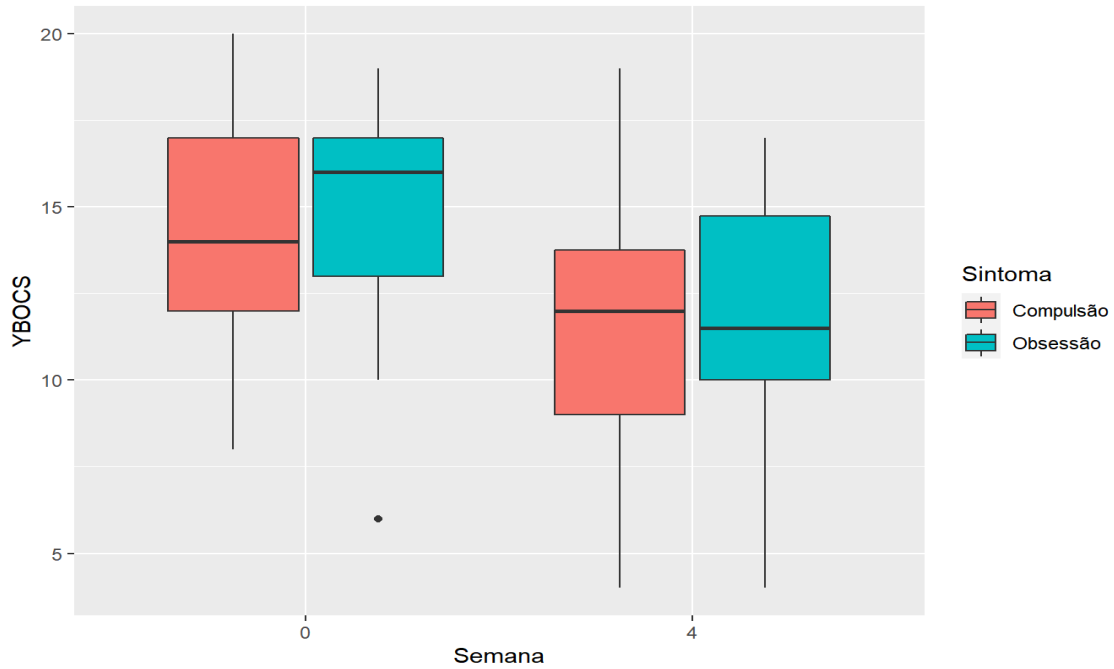


Figura 1: Box plot do YBOCS por semana e por sintomas do TOC

Para que seja possível quantificar a melhora do paciente de uma semana para outra, foi decidido adotar como variável resposta do estudo o percentual de redução do YBOCS da semana 0 para a semana 4, de acordo com a fórmula a seguir:

$$Y_i = \frac{YBOCS_{i,0} - YBOCS_{i,4}}{YBOCS_{i,0}},$$

em que 0 e 4 fazem referência ao número da semana e i é o número do paciente.

A Figura 2 apresenta o histograma da variável resposta para cada sintoma de TOC, sendo que valores negativos indicam uma piora do indivíduo ao longo do estudo e ocorreram duas vezes para o sintoma de compulsão e uma para obsessão.

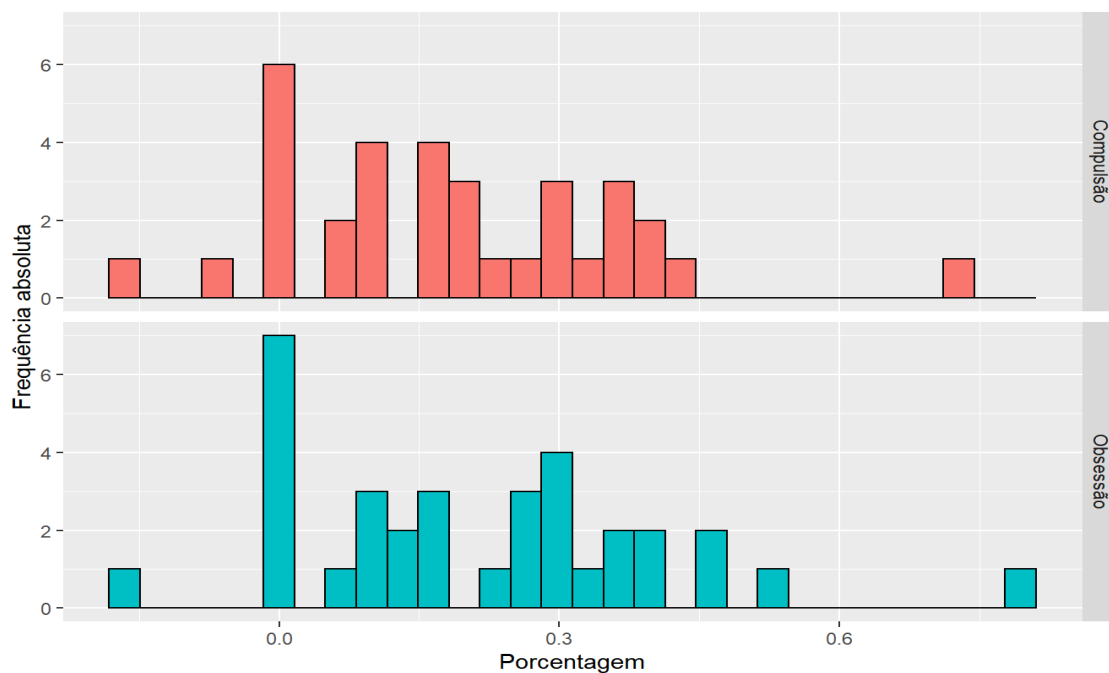


Figura 2: Histograma da variável resposta por sintoma de TOC

5.2. Seleção de variáveis

Como o número de variáveis do estudo (67) é muito superior ao número de indivíduos (37), é um desafio selecionar quais variáveis serão candidatas a marcadores e preditores da melhora do paciente e, para isso, foi utilizado o *Informative Value* (IV) das variáveis (Thomas et al, 2002).

O IV é uma medida muito utilizada para seleção de variáveis no contexto de credit score, ou seja, em modelos cuja a variável resposta é categórica. Dessa forma, categorizou-se a melhora dos pacientes em dois níveis, aqueles que obtiveram redução do YBOCS superior a 20%, considerados como pacientes com uma melhora efetiva e aqueles com redução inferior a 20%, considerados sem melhora efetiva.

A Figura 3 contém os IV's para cada variável por sintoma e, como é comum no uso dessa medida, variáveis que apresentarem IV superior a 0.3, pontos azuis, são consideradas boas candidatas para o modelo (Lund e Brotherton, 2013).

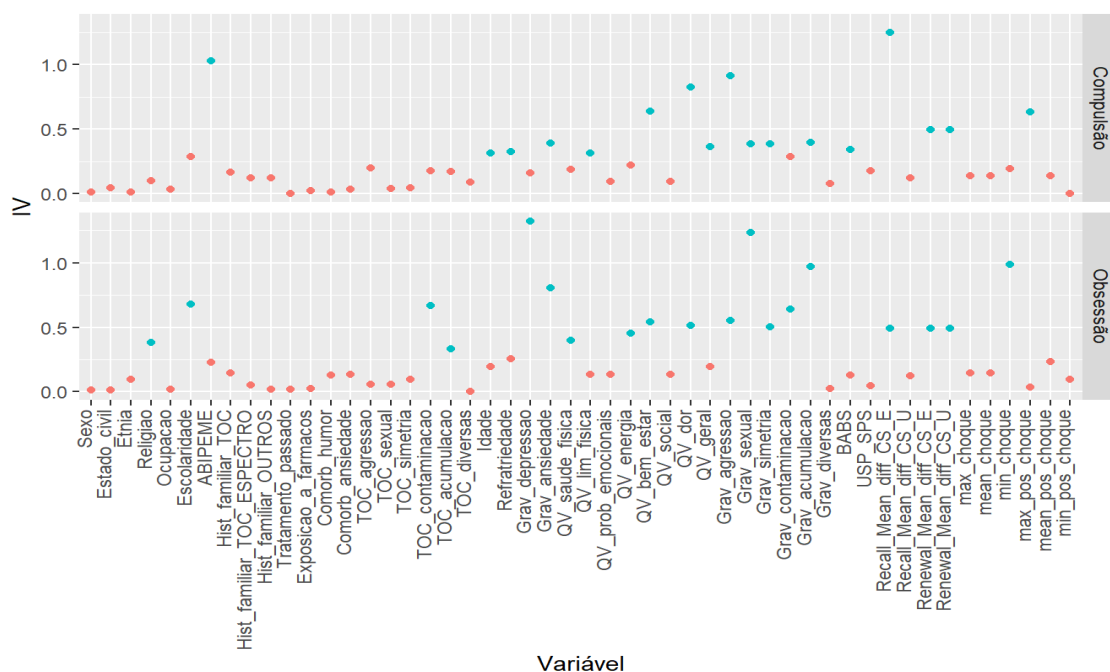


Figura 3: IV das variáveis por sintoma de TOC

A Tabela 2 contém as variáveis que foram selecionadas para cada sintoma e seus respectivos IV's. Através dela percebemos que apesar de terem muitas variáveis em comum, há diferenças entre as variáveis discriminantes para os sintomas, como por exemplo o mínimo do choque se mostra discriminante para a obsessão, mas não para a compulsão e o máximo do pós choque apresenta comportamento inverso.

Tabela 2: Variáveis selecionadas como possíveis preditores e marcadores

Obsessão		Compulsão	
Variável	IV	Variável	IV
Gravidade de depressão de Beck	1,33	Média da diferença de CS_E na recuperação	1,25
Gravidade de sintomas do tipo sexual	1,24	ABIPEME (Classe social)	1,03
Mínimo dos choques	0,99	Gravidade de sintomas do tipo agressão	0,92
Gravidade de sintomas do tipo acumulação	0,97	Qualidade de vida - dor	0,83
Gravidade de sintomas do tipo ansiedade	0,81	Qualidade de vida - bem estar	0,64
Escolaridade	0,68	Máximo das medições do Pós Choque	0,63
Presença de TOC de contaminação	0,67	Média da diferença de CS_E na renovação	0,49
Gravidade de sintomas do tipo contaminação	0,64	Média da diferença de CS_U na renovação	0,49
Gravidade de sintomas do tipo agressão	0,55	Gravidade de sintomas do tipo acumulação	0,40
Qualidade de vida - bem estar	0,54	Gravidade de sintomas do tipo ansiedade	0,39
Qualidade de vida - dor	0,51	Gravidade de sintomas do tipo sexual	0,38
Gravidade de sintomas do tipo simetria	0,50	Gravidade de sintomas do tipo simetria	0,38
Média da diferença de CS_E na recuperação	0,49	Qualidade de vida - geral	0,36
Média da diferença de CS_E na renovação	0,49	BABS	0,34
Média da diferença de CS_U na renovação	0,49	Refratariedade	0,32
Qualidade de vida - energia	0,45	Idade	0,31
Qualidade de vida - saúde física	0,40	Qualidade de vida - lim física	0,31
Religião	0,38	-	-
Presença de TOC de acumulação	0,33	-	-

5.3. Análise descritiva das variáveis selecionadas

Conforme alinhado com os pesquisadores, serão ajustados modelos separados para obsessões e compulsões, portanto a análise descritiva será feita da mesma forma, separando entre variáveis categóricas e numéricas por questão de organização. É importante salientar também que, pelo elevado número de variáveis, seria inviável e massante que todos os gráficos fossem incluídos no corpo principal do relatório, podendo ser encontrados integralmente na pasta “GRAFICOS” na documentação final do projeto.

Para compulsões, de acordo com o critério de valor informativo, foram selecionadas 17 variáveis, sendo apenas uma delas categórica (classe social), cujo box plot está apresentado na Figura 4.

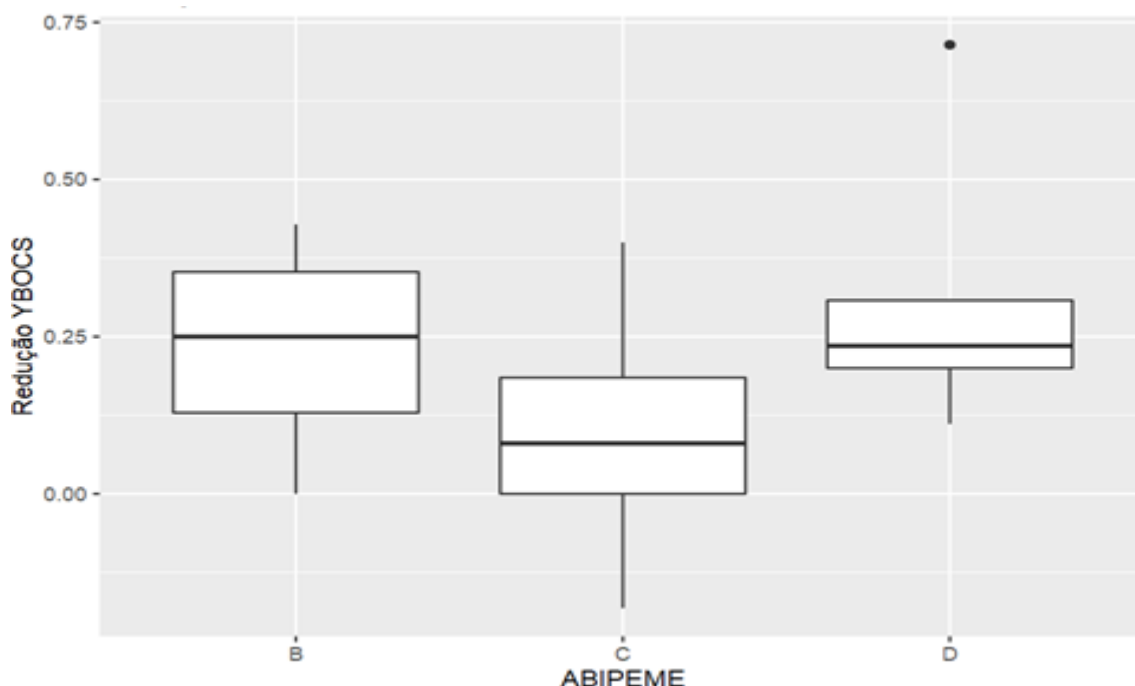


Figura 4: Box plot da redução de YBOCS por classe social.

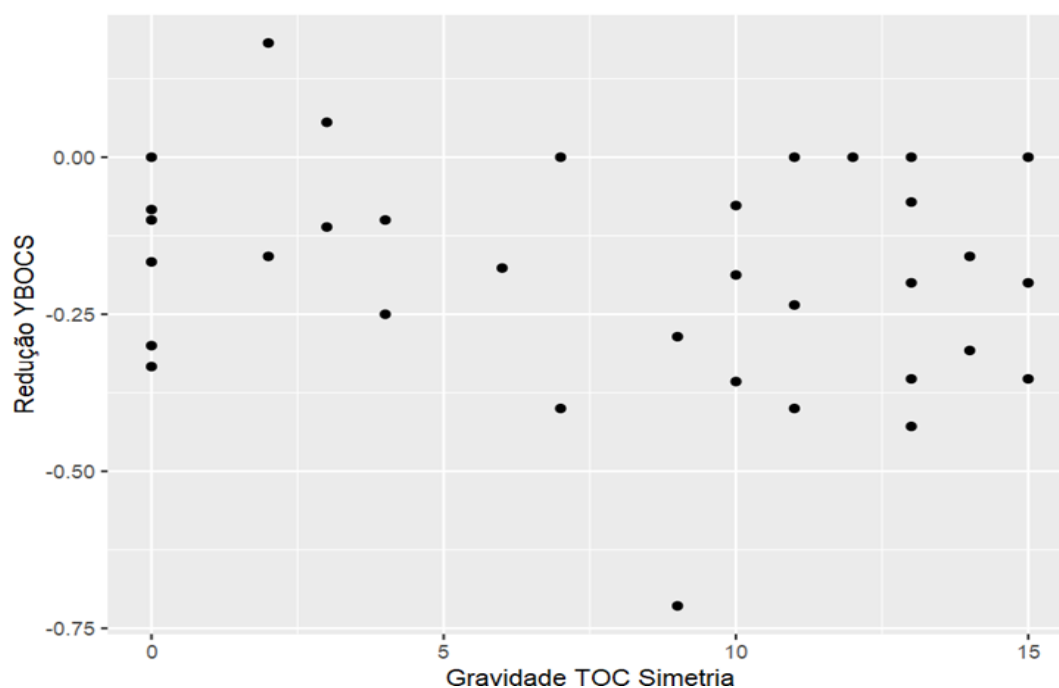
Nota-se primeiramente que a classe social C se destaca com uma mediana mais próxima de zero do que as outras classes, ou seja, os indivíduos pertencentes apresentam, em mediana, redução no YBOCS menor que as outras classes envolvidas. Há também semelhança entre as medianas das classes D e B, mas é

necessária cautela ao fazer a comparação, pois o tamanho amostral entre elas não é uniforme.

Dentre as variáveis numéricas, os seguintes grupos podem ser extraídos:

- Variáveis de base
- Variáveis de choque
- Variáveis de renovação/recuperação.

Do primeiro tópico destacam-se as variáveis de gravidade dos espectros de transtorno, como TOC do tipo simetria e TOC do tipo de agressão, cujas dispersões em relação à redução do YBOCS constam a seguir.



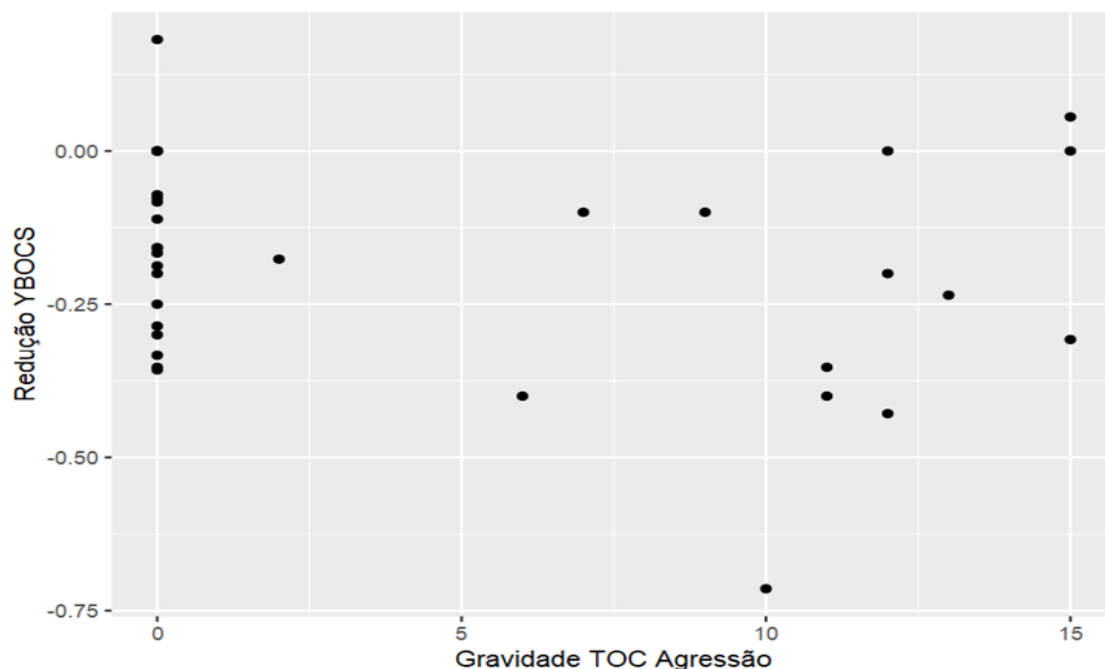


Figura 5 : Dispersão da redução de YBOCS em função de gravidade do TOC tipos simetria e agressão.

O primeiro ponto a destacar são os pacientes que não apresentam o TOC dos tipos agressão e simetria, constando com valor zero. Para simetria, são menos numerosos os casos zerados, sendo possível notar um aumento na redução dos níveis YBOCS quando a gravidade do TOC tipo simetria está entre 5 e 10. Já para o tipo agressão o número de observações nulas é muito maior, com apenas 14 pontos diferentes de zero. Para níveis mais altos de gravidade do TOC agressão a linha de tendência apresenta uma leve queda, entretanto os dados ainda estão em um intervalo muito similar aos pacientes que não possuem esse tipo.

Para as variáveis de choque (e pós-choque), foi alinhado com os pesquisadores que exploraremos três medidas resumo dos 10 choques, sendo elas o máximo, o mínimo e a média. Para compulsão, o IV foi suficiente para o máximo do pós-choque, cujo gráfico com linha de tendência consta na Figura 6.

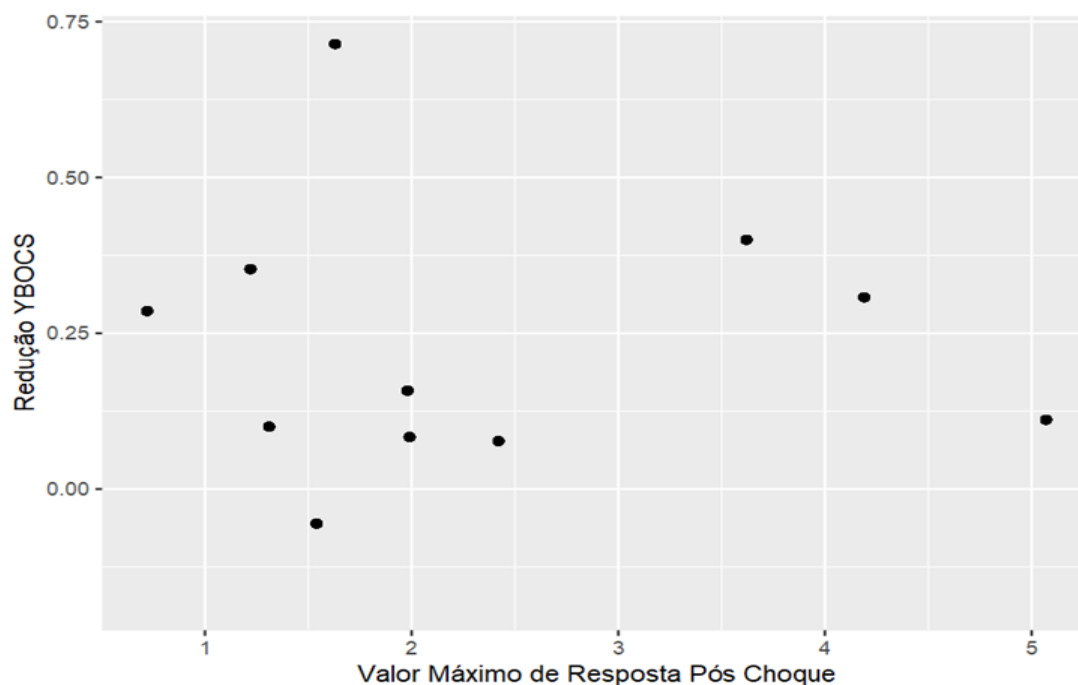
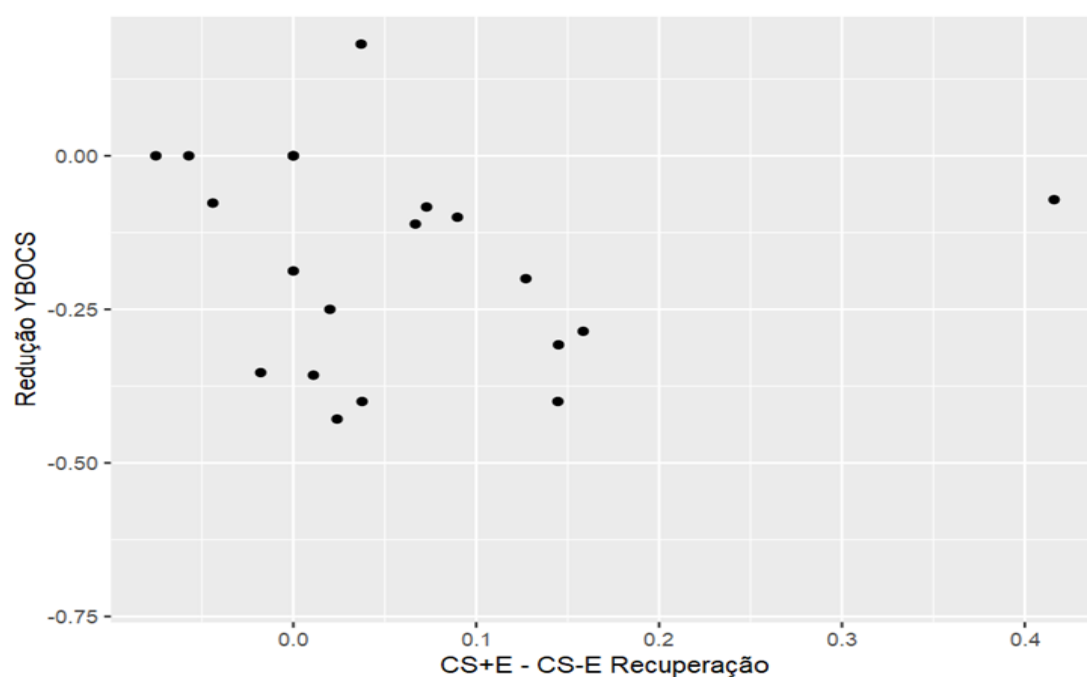
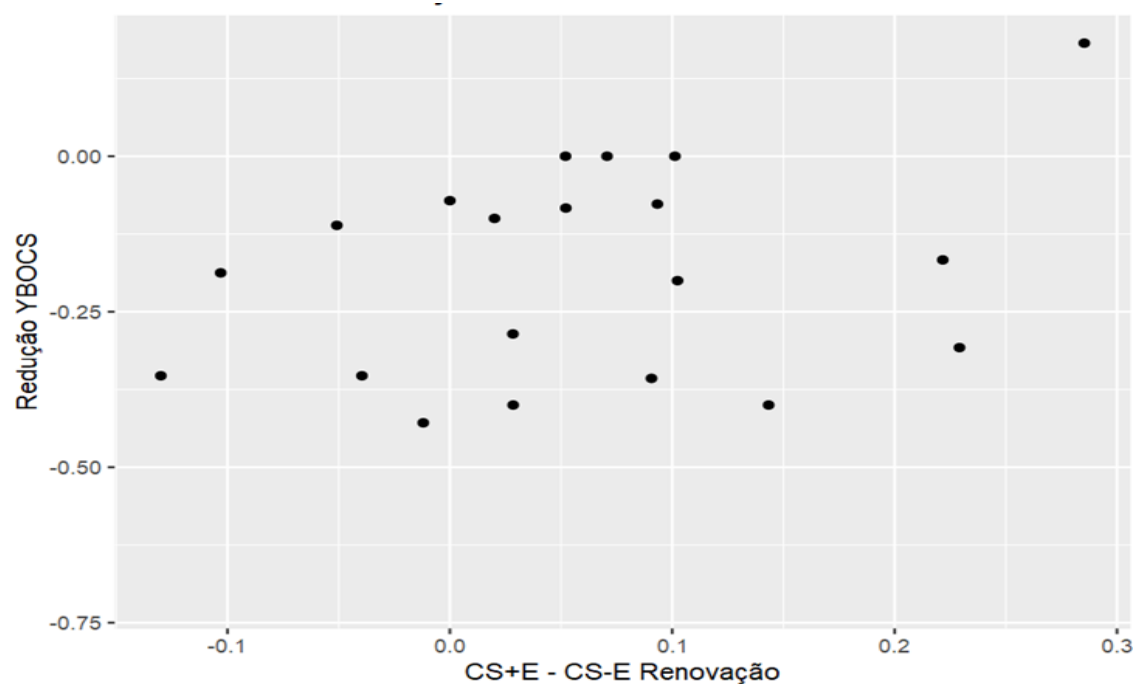


Figura 6: Redução de YBOCS em função do máximo de pós-choque.

Quando são comparados o número de observações das variáveis choque e pós-choque há uma diferença grande, sendo que no pós choque há apenas 11 observações no gráfico. Por esse motivo, a tendência que a suavização indica pode não ser condizente com o comportamento real dos dados.

Têm-se finalmente as variáveis de renovação e recuperação, ou seja, as médias das diferenças entre os estímulos extintos com e sem choque, e não extintos com e sem choque, na recuperação e renovação.



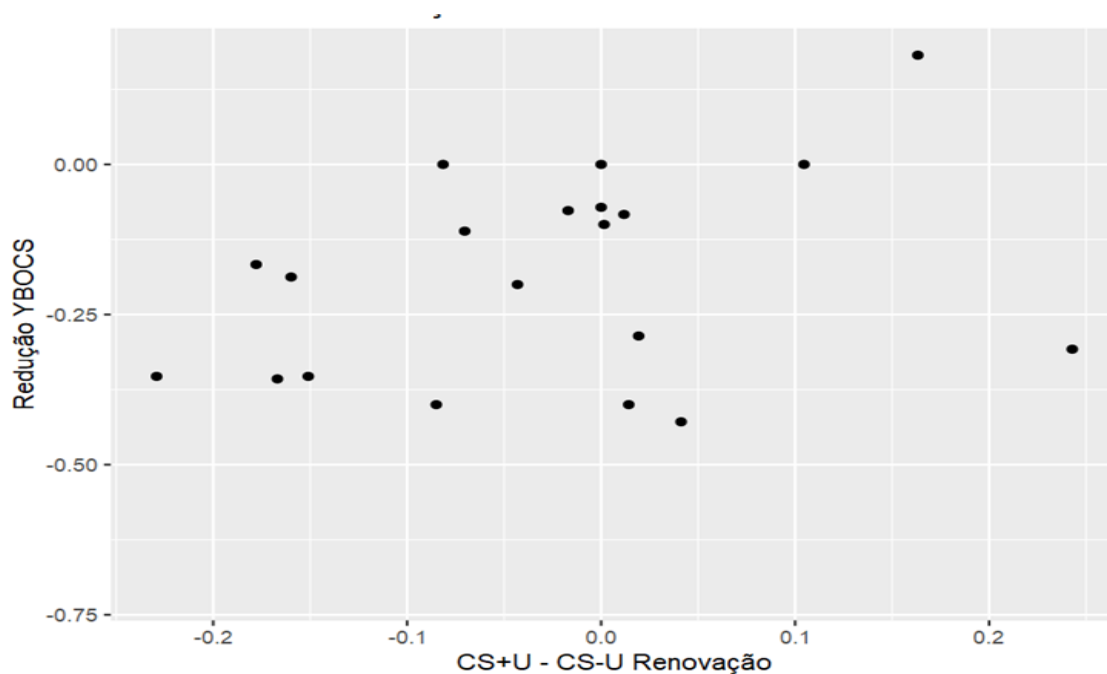


Figura 7: Gráficos de dispersão da diminuição de YBOCS de compulsão por caso de renovação ou recuperação.

Nota-se que as diferenças CS+U - CS-U e CS+E - CS-E possuem tendências mais definidas, a primeira indicando que conforme aumenta a diferença, menor tende a ser a redução de YBOCS quando não há a extinção, e a segunda indicando que quanto maior a diferença, maior tende a ser a redução de YBOCS, quando há o processo de extinção do estímulo. Para CS+E - CS-E na recuperação, a tendência seria de aumento, e destaca-se a última observação que puxa a tendência para baixo.

Para a medida de obsessão, a análise foi análoga e está apresentada nas Figuras 8 a 12.

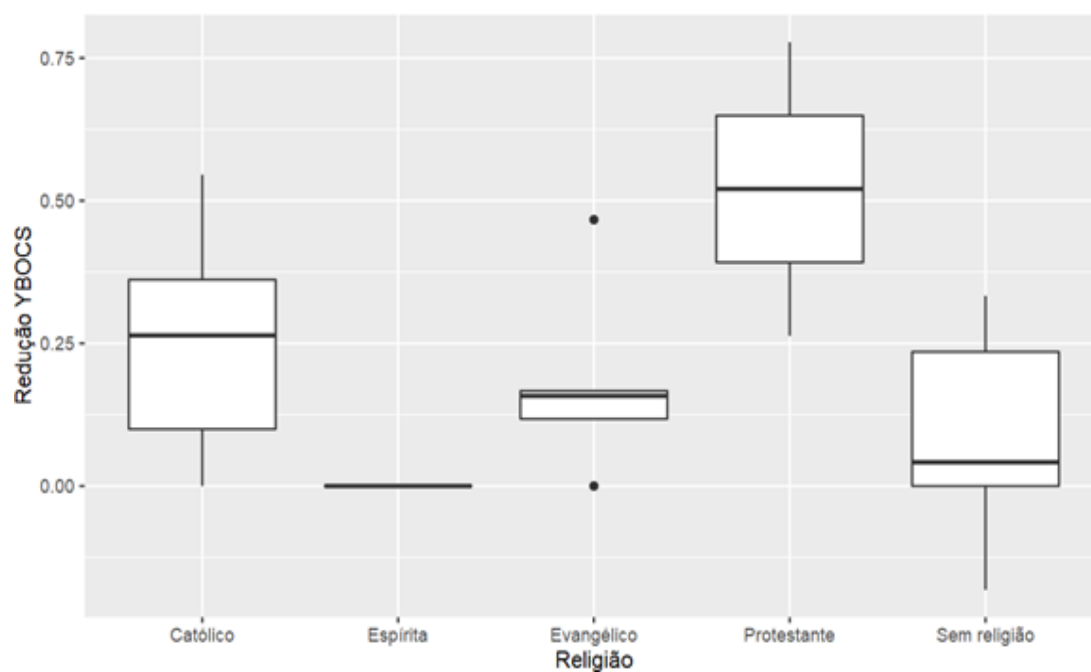
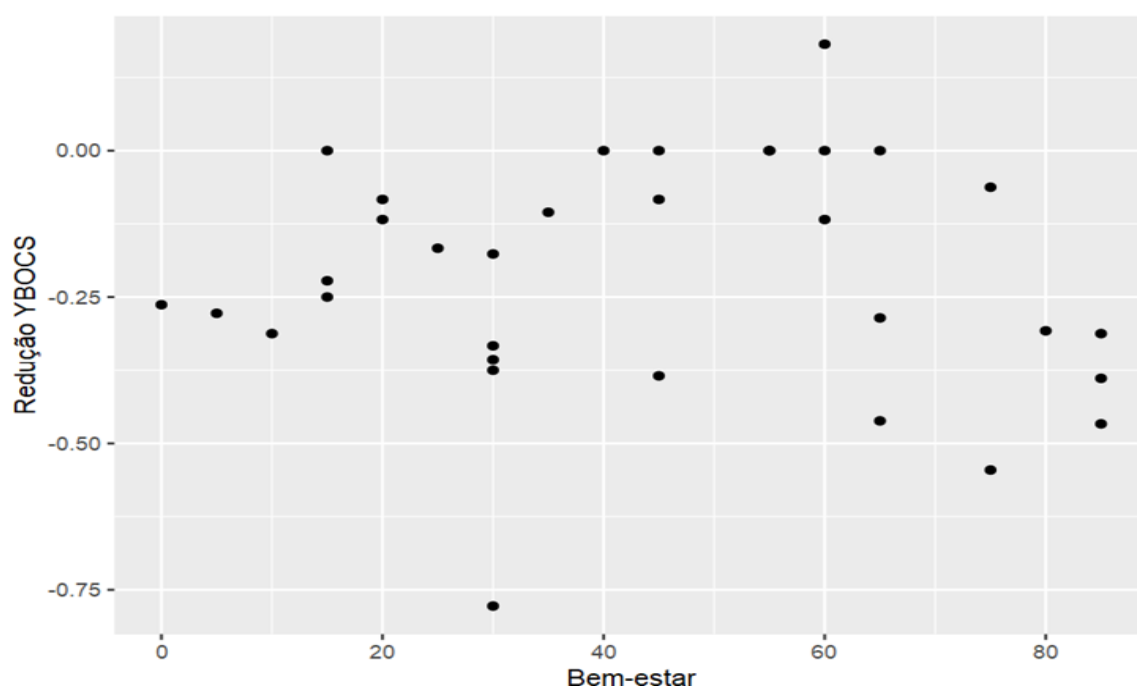


Figura 8: Box plots da redução de YBOCS de obsessão por nível de religião.



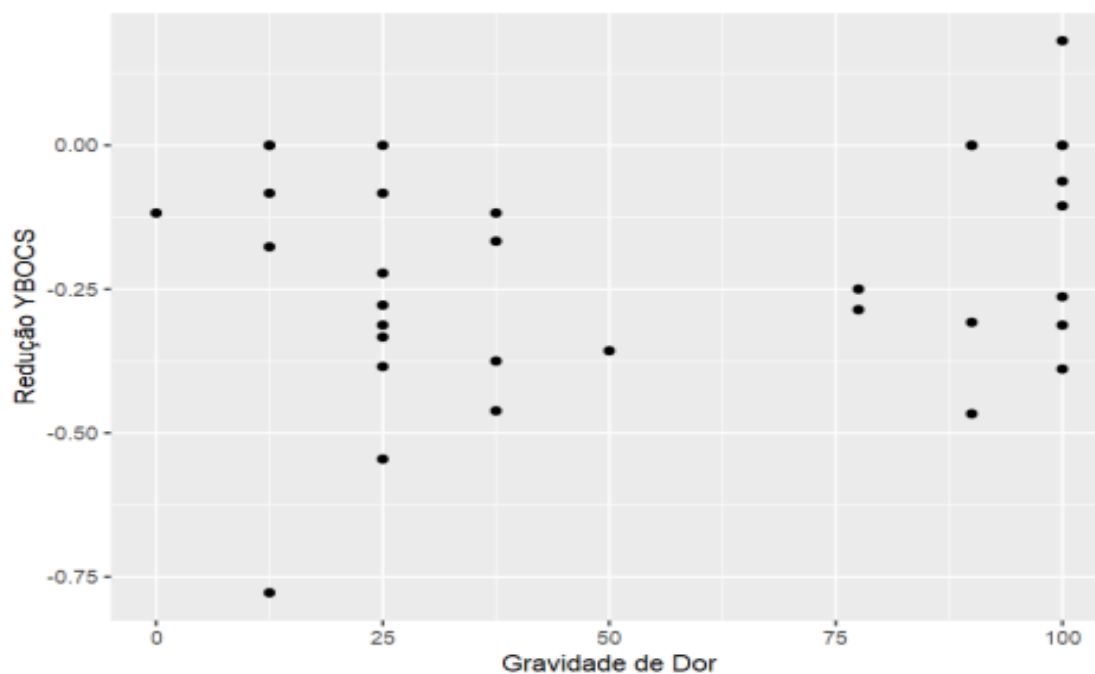


Figura 9: Dispersão da diminuição de YBOCS de obsessão por escala de bem-estar e de gravidade de dor.

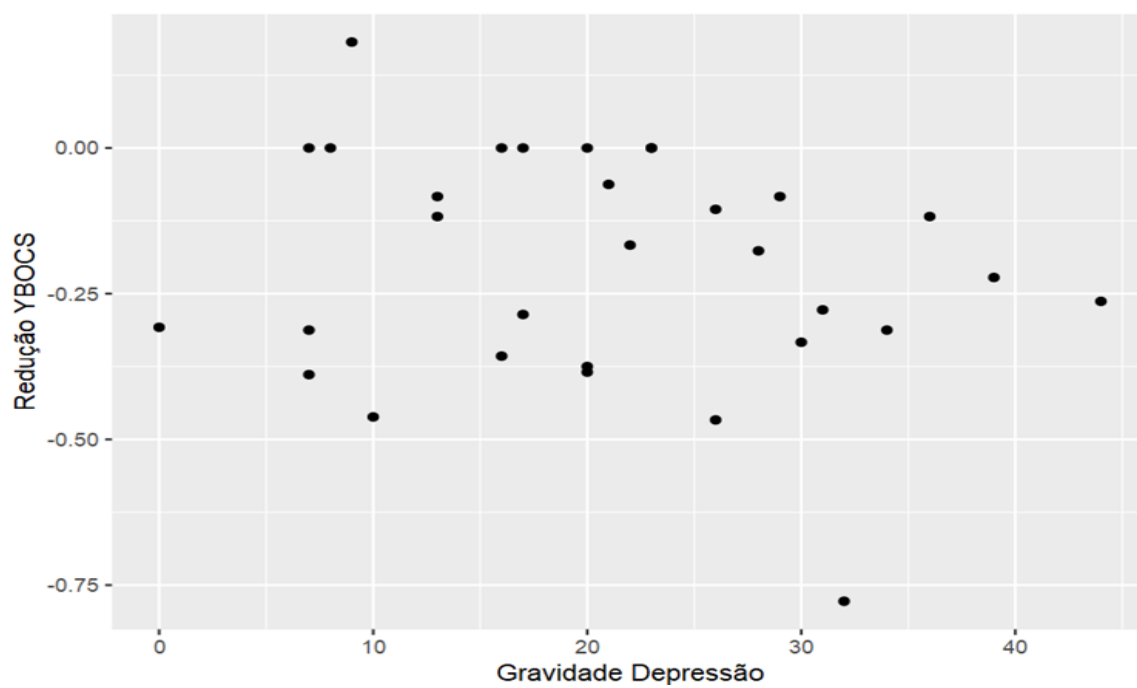


Figura 10: Redução de YBOCS de obsessão por gravidade de depressão de Beck.

Para religião o comentário sobre o tamanho amostral dos grupos também é válido, uma vez que 29 dos 37 indivíduos que compõem o banco de dados declaram-se como católicos, evangélicos ou protestantes. Entre os três grupos de maior evidência, há boa diferenciação entre as medianas dos protestantes e dos outros grupos. Embora a caixa dos evangélicos esteja contida no mesmo intervalo que os católicos, a amplitude é muito menor, podendo indicar diferenciação.

Para a medida de bem-estar a curva de tendência evidencia uma mudança visível na redução de YBOCS a partir da faixa de 40, mas depois aumentando em magnitude maior, o que faria sentido uma vez que espera-se intuitivamente que aumento de bem-estar implique na redução da escala de TOC. Já para a escala de dor, fora o primeiro ponto com gravidade nula nota-se tendência de menores diminuições em YBOCS conforme a dor aumenta, o que intuitivamente também é aceitável.

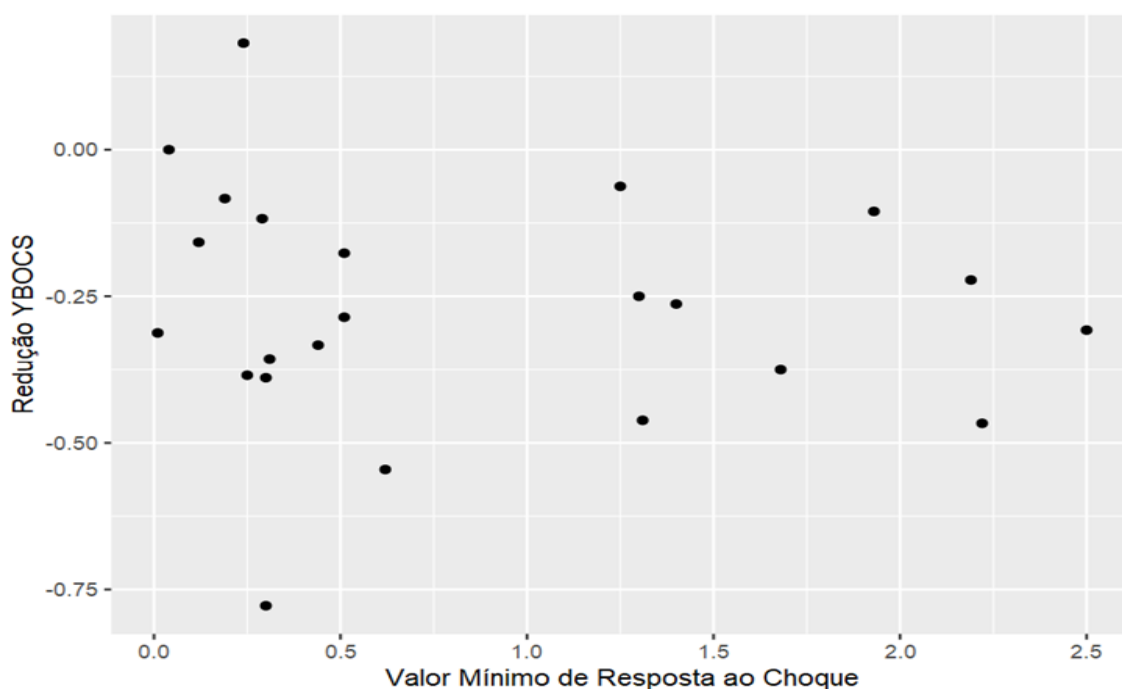
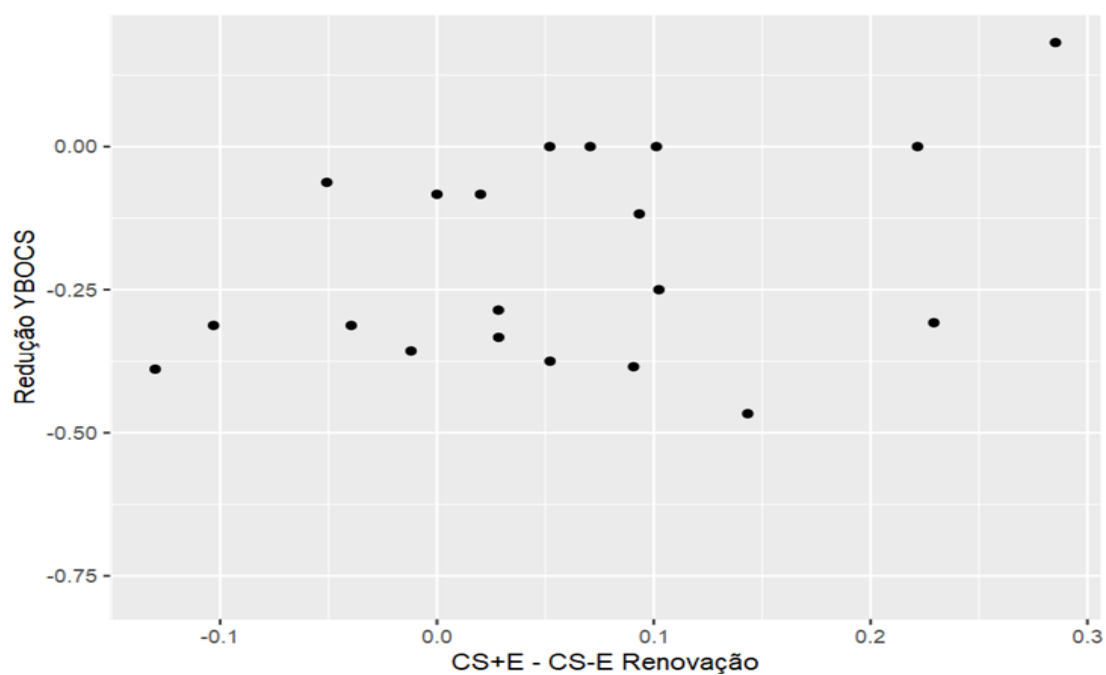
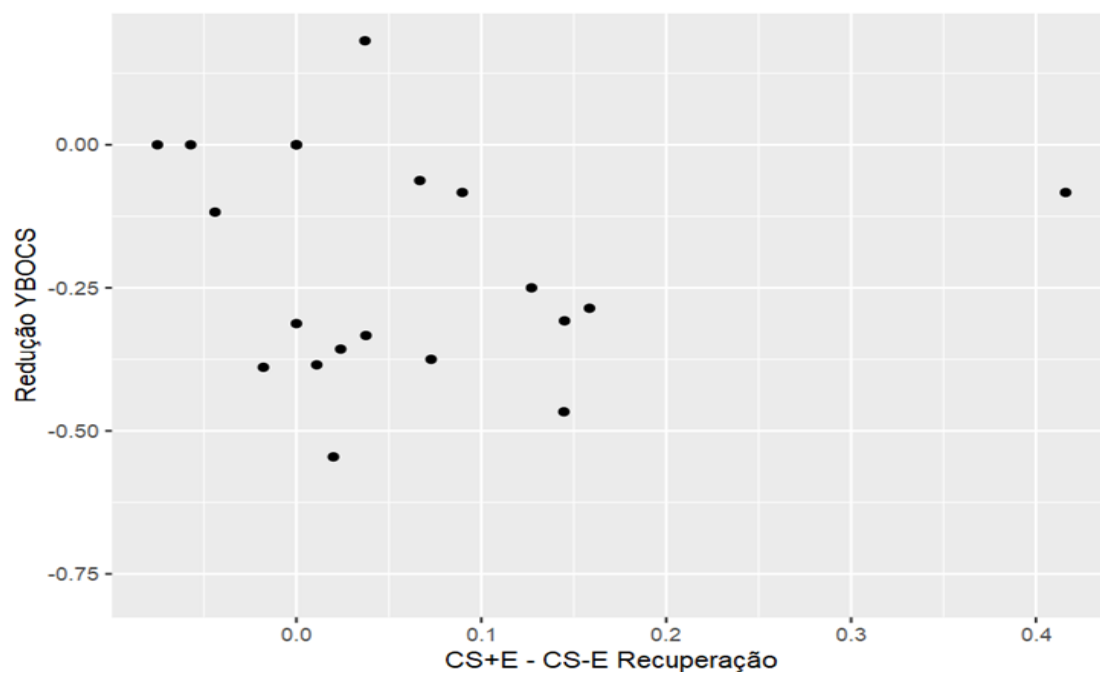


Figura 11: Dispersão da redução de YBOCS de obsessão em função do valor mínimo de resposta ao choque.

Com relação a variáveis de choque, percebe-se inicialmente um aumento acentuado conforme o mínimo de resposta observada fica entre 0 e 0.75, tendência que parece aproximadamente estável após esse valor, podendo indicar um limiar.



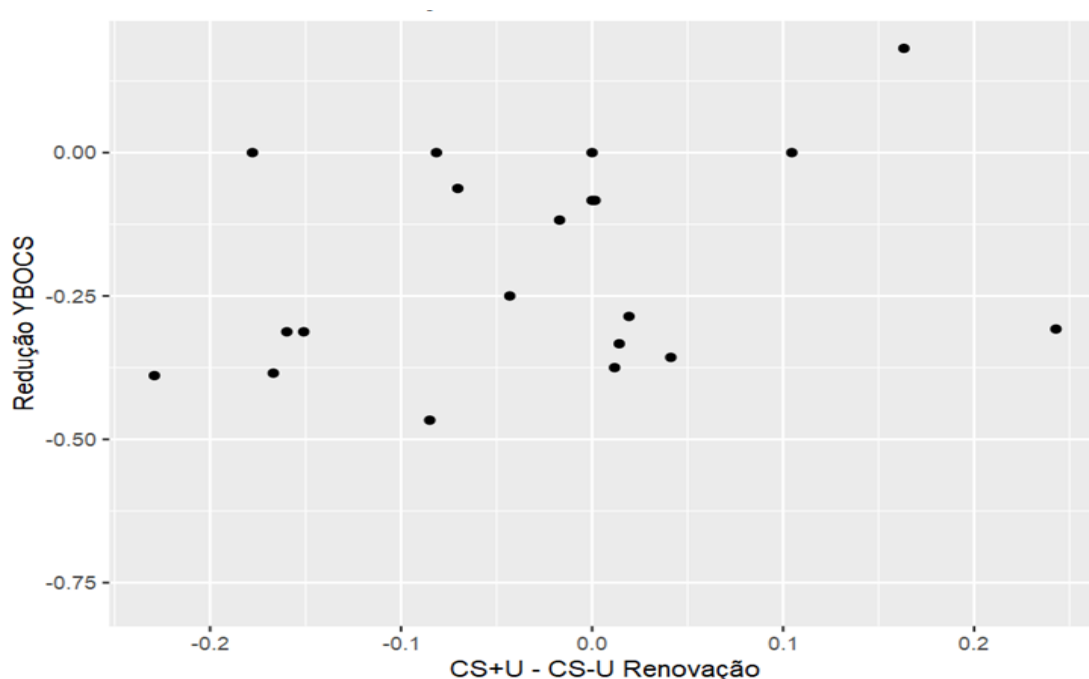


Figura 12: Dispersões da diminuição de YBOCS de obsessão por caso de renovação ou recuperação.

Considerando variáveis de renovação, tanto para o processo com extinção quanto para sem extinção nota-se uma tendência de diminuição na redução de YBOCS de obsessão conforme aumenta a diferença entre as situações com e sem choque. Na recuperação, o processo com extinção de estímulo mostra tendência contrária, ou seja, aumento da diminuição de YBOCS conforme aumenta a diferença entre as situações com e sem choque.

O próximo passo agora é realizar análises que possibilitem identificar marcadores e preditores da redução na escala YBOCS, ou da melhora dos pacientes (definida pelos pesquisadores como redução maior ou igual a 20% na escala), sendo a próxima sessão dedicada a essas análises.

6. Análise Inferencial

A análise inferencial será particionada, primeiramente sob a ótica de variável numérica (redução na escala YBOCS), e posteriormente como variável categórica (houve melhora ou não). Como subdivisões, serão feitas análises dos sintomas de obsessão e compulsão para os possíveis marcadores e preditores de melhora ou redução da escala YBOCS.

A quantidade de observações em cada subdivisão variou devido as informações faltantes nas variáveis selecionadas como preditores e marcadores, dessa forma, foi-se utilizado diferentes níveis descritivos para seleção de variáveis nos modelos de regressão, seguindo Pericchi e Pereira (2016) e considerando uma amostra ideal de 50 indivíduos. Os níveis descritivos adaptados estão na Tabela 3.

Tabela 3: valores p adaptados em função da amostra

Subdivisão	n	alfa de 5%	alfa de 10%
preditores numéricas	28	7,2%	14,4%
preditores categóricas	28	7,2%	14,4%
marcadores numéricos	18	9,4%	19,0%
marcadores categóricas	13	11,6%	23,5%

6.1. Variável Numérica

O método escolhido para a abordagem numérica foi a regressão linear simples, e para cada modelo (preditores de redução para obsessão e compulsão, e marcadores de redução para obsessão e compulsão) serão apresentados as variáveis do modelo final, nível de significância das variáveis selecionadas, análise diagnóstica e interpretação dos parâmetros.

- **6.1.1. Regressão linear para obsessão**

- Preditores

Foram utilizadas inicialmente as variáveis selecionadas listadas via informative value, listadas na Tabela 2, com exceção das que são medidas em momentos posteriores, caso de pós-choque e variáveis de renovação e recuperação, que são candidatas a marcadores. A pedido dos pesquisadores, foram consideradas em todas as análises para preditores o sexo e a idade dos pacientes.

Dada a baixa incidência de alguns níveis das variáveis Escolaridade e Religião, foram feitas algumas aglutinações: protestantes foram unidos aos católicos, e os espíritas (apenas uma observação) foram unidos ao grupo sem religião. Já aqueles que eram analfabetos, foram unidos aos que não possuíam ensino fundamental completo.

A Tabela 4 mostra as variáveis inicialmente consideradas, e seus valores p.

Tabela 4: Modelo completo de regressão linear para preditores de obsessão

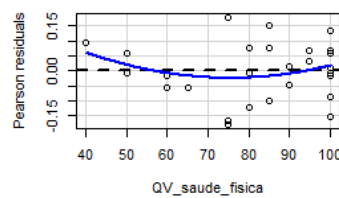
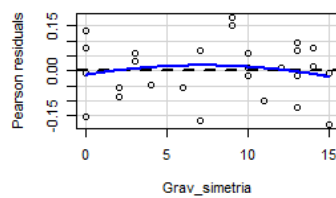
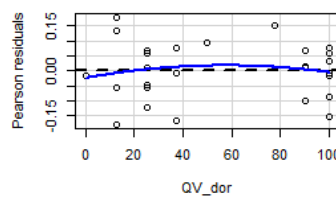
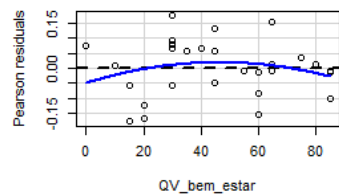
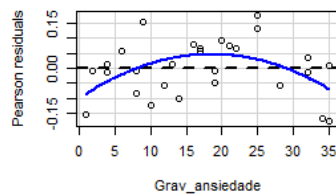
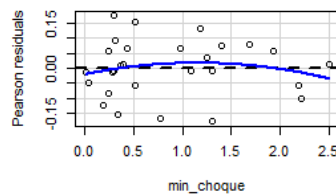
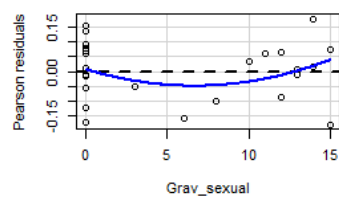
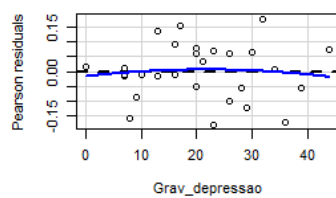
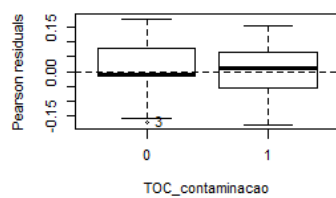
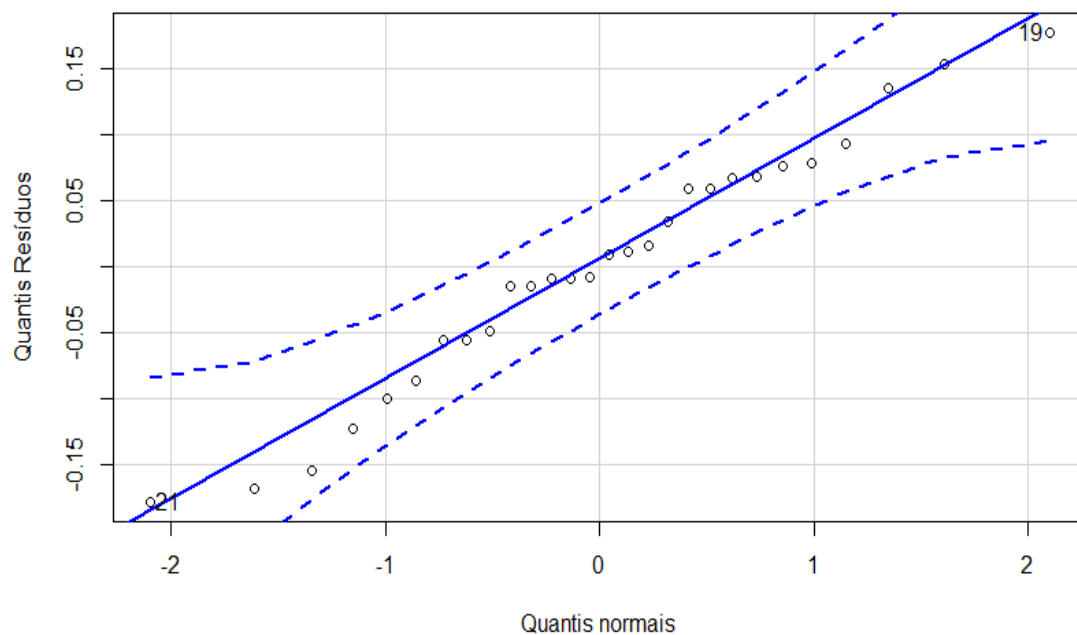
Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,559	0,747	0,749	0,482
Sexo2	-0,042	0,091	-0,463	0,659
Idade	-0,002	0,010	-0,209	0,841
Escolaridade5	-0,023	0,279	-0,081	0,938
Escolaridade6	0,202	0,237	0,853	0,427
Escolaridade7	-0,027	0,210	-0,130	0,901
Escolaridade8	-0,058	0,385	-0,149	0,886
TOC_contaminacao1	-0,646	0,419	-1,542	0,174
Grav_depressao	0,011	0,010	1,165	0,288
Grav_sexual	0,008	0,007	1,074	0,324
min_choque	0,115	0,060	1,900	0,106
Grav_acumulacao	-0,025	0,064	-0,392	0,709
Grav_ansiedade	-0,015	0,007	-2,253	0,065
Grav_contaminacao	0,030	0,031	0,987	0,362
Grav_agressao	-0,003	0,009	-0,293	0,779
QV_bem_estar	0,006	0,004	1,668	0,146
QV_dor	-0,004	0,002	-2,237	0,067
Grav_simetria	0,007	0,010	0,695	0,513
QV_energia	-0,001	0,003	-0,396	0,706
QV_saude_fisica	-0,004	0,005	-0,730	0,493
Religiao7	0,006	0,212	0,028	0,979
TOC_acumulacao1	0,255	0,391	0,654	0,538

A partir desse modelo, foi utilizado o método de seleção de Akaike para realizar um stepwise em ambos sentidos, o que resultou no modelo com as variáveis apresentadas na Tabela 5.

Tabela 5: Modelo final de regressão linear para preditores de obsessão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,131	0,172	0,764	0,455
TOC_contaminacao1	-0,269	0,049	-5,517	0,000
Grav_depressao	0,014	0,004	3,949	0,001
Grav_sexual	0,010	0,004	2,719	0,014
min_choque	0,097	0,035	2,745	0,013
Grav_ansiedade	-0,010	0,003	-3,203	0,005
QV_bem_estar	0,007	0,002	4,026	0,001
QV_dor	-0,004	0,001	-3,991	0,001
Grav_simetria	0,011	0,004	2,599	0,018
QV_saude_fisica	-0,002	0,001	-1,523	0,145

Deve-se lembrar que, pelo baixo tamanho amostral, utilizamos a relação já citada com o nível de significância. Na Figura 13, é demonstrado o bom ajuste do modelo através da análise diagnóstica.



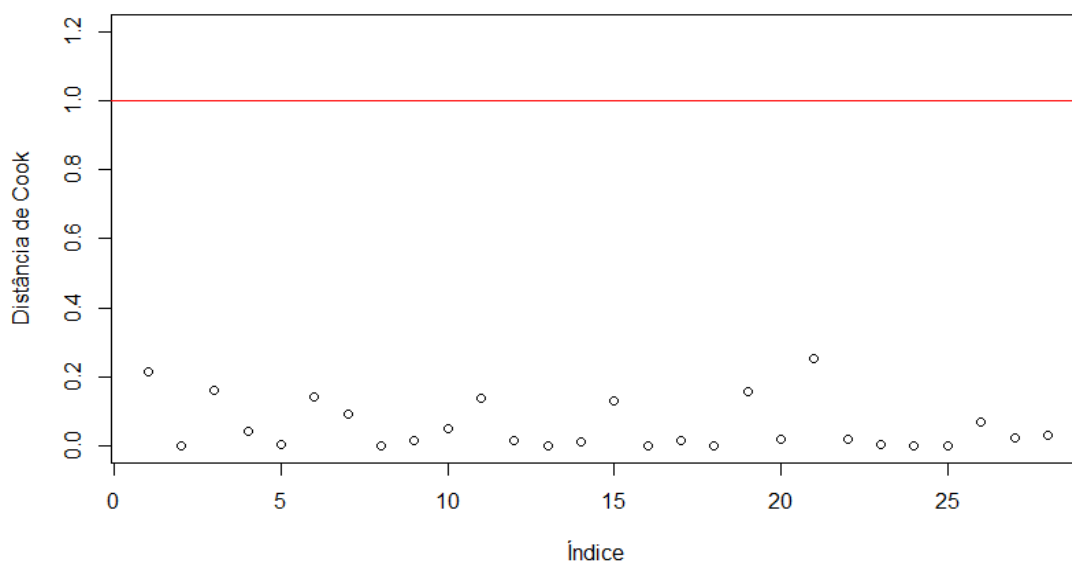


Figura 13: Gráfico QQ, resíduos por variável, e distância de Cook - modelo final (Tabela 5).

Os gráficos acima demonstram o bom ajuste do modelo. Os resíduos estão dentro da banda de 95% de confiança contra os quantis normais, indicando que a premissa de normalidade dos resíduos é satisfeita. Nota-se também que os resíduos se comportam de modo aceitável ao redor do zero, sem apresentar grandes tendências apesar do pequeno tamanho amostral, dispersos em uma amplitude que parece constante, indicando que são satisfeitas as suposições de média zero e variância constante. Já as distâncias de Cook descartam a possibilidade de que um ponto seja influente em excesso.

Chega-se então à conclusão de que a presença de TOC de tipo contaminação, a gravidade dos sintomas de depressão, a gravidade do TOC tipo sexual, o mínimo das medidas de choque, a gravidade dos sintomas de ansiedade, a gravidade do TOC tipo simetria, e os scores de bem-estar, dor e saúde física, são conjuntamente bons candidatos a preditores de redução da escala YBOCS, para obsessão.

A interpretação dos coeficientes que constam na Tabela 5 é a seguinte:

- TOC_Contaminação - a presença de TOC tipo contaminação diminui o valor esperado da redução da escala YBOCS em 0,269 em relação a uma pessoa

que não possui o TOC tipo contaminação, mantidas todas as outras variáveis constantes.

- Grav_depressao - um aumento de um ponto na escala de gravidade de depressão causa um incremento de 0,014 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- Grav_sexual - um aumento de um ponto na escala de gravidade de TOC tipo sexual causa um incremento de 0,010 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- min_choque - o aumento de uma unidade no valor mínimo das variáveis de choque causa um incremento de 0,097 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- Grav_ansiedade - um aumento de um ponto na escala de gravidade de ansiedade causa um decréscimo de 0,010 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- Grav_simetria - um aumento de um ponto na escala de gravidade de TOC tipo simetria causa um incremento de 0,011 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- QV_bem_estar - um aumento de um ponto na escala de qualidade de vida ligada ao bem-estar do paciente causa um incremento de 0,007 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- QV_dor - um aumento de um ponto na escala de qualidade de vida ligada à dor que o paciente sente causa um decréscimo de 0,004 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.
- QV_saude_fisica - um aumento de um ponto na escala de qualidade de vida ligada à saúde física do paciente causa um decréscimo de 0,002 na redução esperada da escala YBOCS, mantidas todas as outras variáveis constantes.

Nota-se que a presença do TOC tipo contaminação aparenta ter um efeito muito forte sob a redução da escala YBOCS para preditores de obsessão.

Os modelos de regressão apresentados a seguir possuem interpretação análoga.

- Marcadores

Mais uma vez, ajustou-se o modelo inicial com todas as variáveis sugeridas, conforme a Tabela 6.

Tabela 6: Modelo completo de regressão linear para marcadores de obsessão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,185	0,065	2,840	0,013
Renewal_Mean_diff_CS_E	-0,135	0,583	-0,231	0,820
Renewal_Mean_diff_CS_U	-0,564	0,530	-1,064	0,305
Recall_Mean_diff_CS_E	0,248	0,410	0,603	0,556

Após redução das variáveis, obteve-se o modelo com as variáveis apresentadas na Tabela 7.

Tabela 7: Modelo reduzido de regressão linear para marcadores de obsessão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,193	0,041	4,662	0,000
Renewal_Mean_diff_CS_U	-0,604	0,369	-1,636	0,121

Embora para o tamanho amostral o nível de significância fosse adequado, e que fossem satisfeitas as condições de normalidade, variância constante e média nula, através da distância de Cook foi encontrada uma observação próxima de ser considerada excessivamente influente.

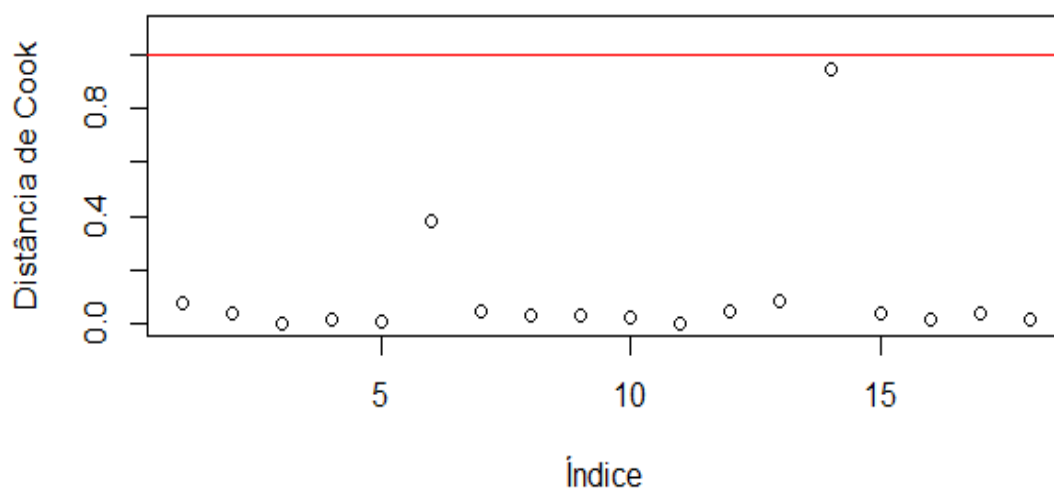


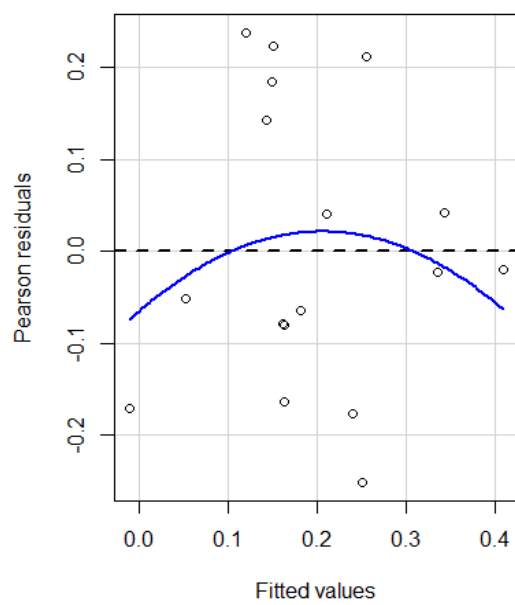
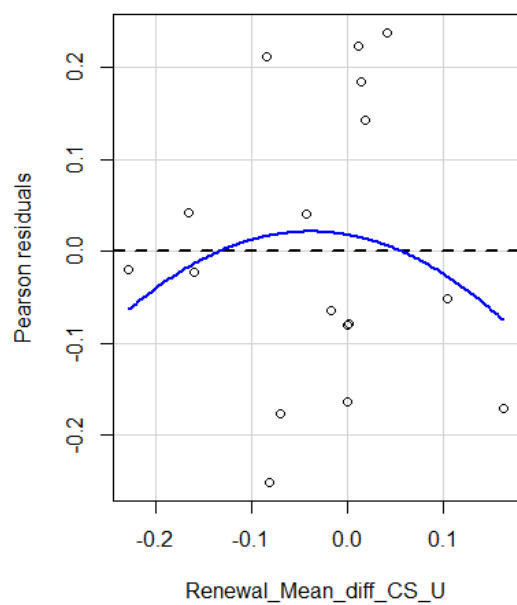
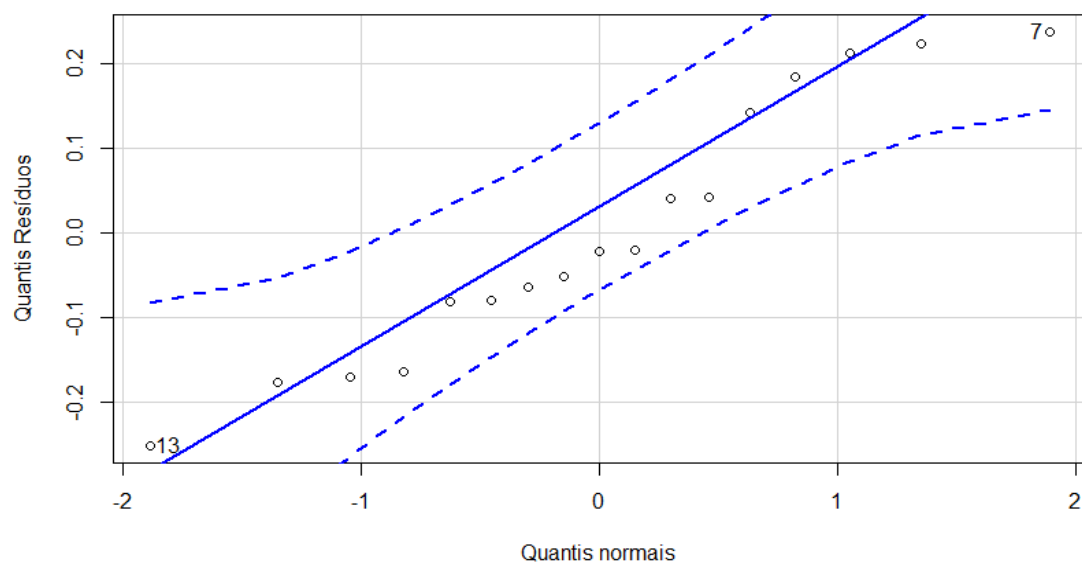
Figura 14: Distância de Cook - modelo reduzido (Tabela 7)

Assim sendo, ajustou-se o modelo sem a observação como modo de avaliar a real influência. Obteve-se a mesma variável como possível marcador, mas mais significativa (Tabela 8).

Tabela 8: Modelo final de regressão linear para marcadores de obsessão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,164	0,040	4,058	0,001
Renewal_Mean_diff_CS_U	-1,071	0,408	-2,628	0,019

Na Figura 16, a análise diagnóstica mostra o bom ajuste.



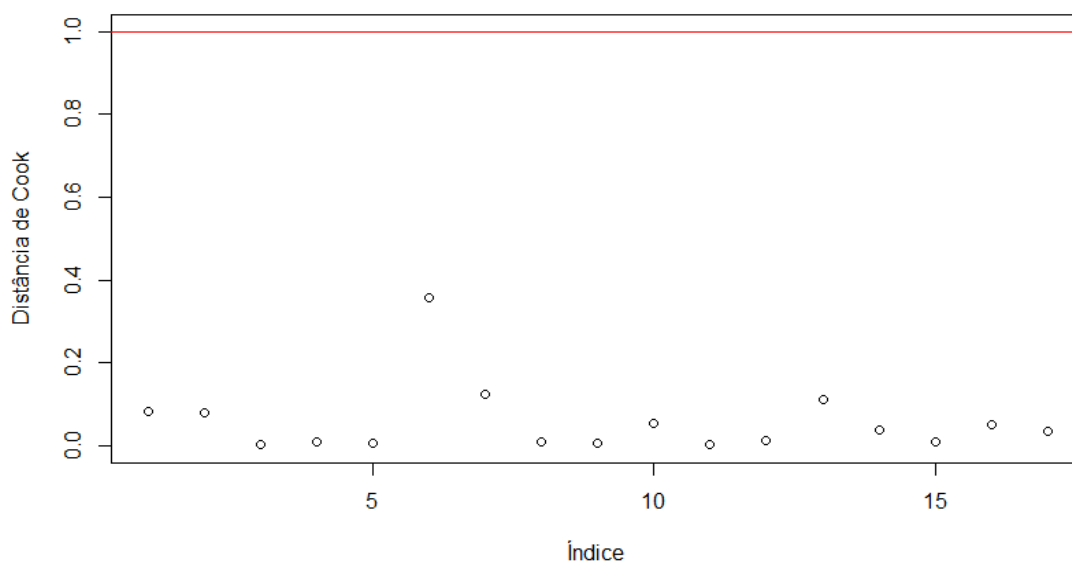


Figura 15: Gráfico QQ, resíduos e distância de Cook - modelo final (Tabela 8)

Os gráficos demonstram que o comportamento dos resíduos está de acordo com as condições de normalidade, homocedasticidade e média zero dos resíduos. Portanto, conclui-se que a variável de renovação sem choque é um bom marcador de redução da escala YBOCS.

• 6.1.2. Regressão linear para compulsão

○ Preditores

Iniciou-se a análise com o modelo completo (Tabela 9):

Tabela 9: Modelo completo de regressão linear para preditores de compulsão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,118	0,399	0,295	0,772
Sexo2	-0,167	0,110	-1,515	0,152
Idade	0,000	0,006	-0,059	0,954
ABIPEME3	-0,095	0,099	-0,966	0,350

Grav_agressao	0,006	0,009	0,618	0,546
QV_dor	-0,002	0,002	-0,748	0,467
QV_bem_estar	-0,003	0,003	-0,931	0,368
Grav_acumulacao	0,011	0,023	0,490	0,632
Grav_ansiedade	-0,003	0,006	-0,516	0,614
Grav_sexual	0,015	0,008	1,825	0,089
Grav_simetria	-0,006	0,010	-0,561	0,584
QV_geral	0,007	0,005	1,569	0,139
BABS	0,008	0,009	0,860	0,404
Refratriedade	-0,020	0,039	-0,507	0,620
QV_lim_fisica	0,000	0,004	-0,074	0,942

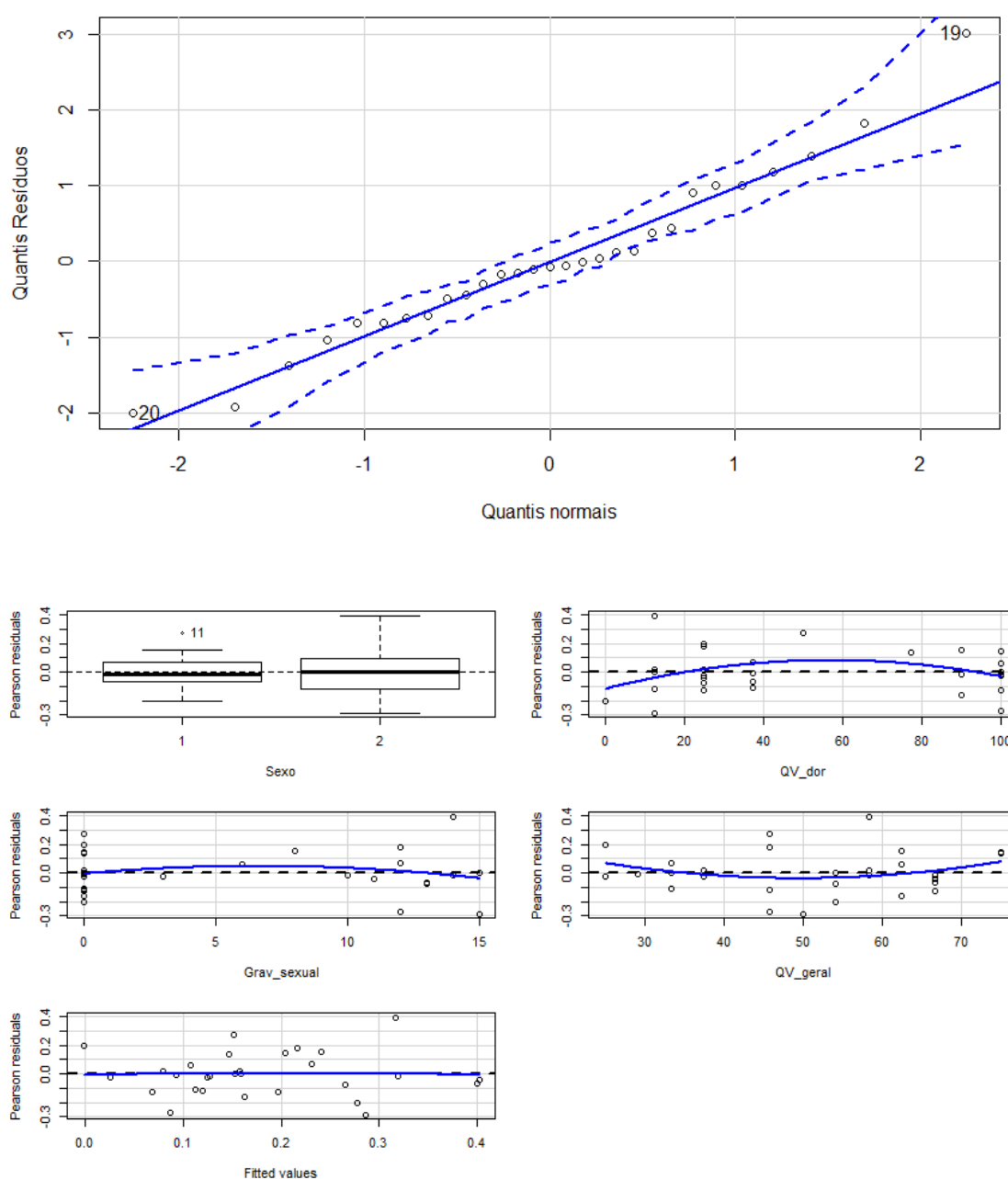
Reduzindo o número de variáveis via stepwise em ambas as direções, obteve-se um modelo com 6 variáveis explicativas. Entretanto, através de redução manual das variáveis foi possível chegar a um modelo com 4 variáveis explicativas, e ajuste satisfatório, apresentado na Tabela 10.

Tabela 10: Modelo final de regressão linear para preditores de compulsão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,018	0,119	0,154	0,879
Sexo2	-0,096	0,065	-1,477	0,153

QV_dor	-0,002	0,001	-1,835	0,079
Grav_sexual	0,010	0,005	1,915	0,068
QV_geral	0,005	0,002	2,091	0,047

A análise diagnóstica do modelo consta na Figura 16.



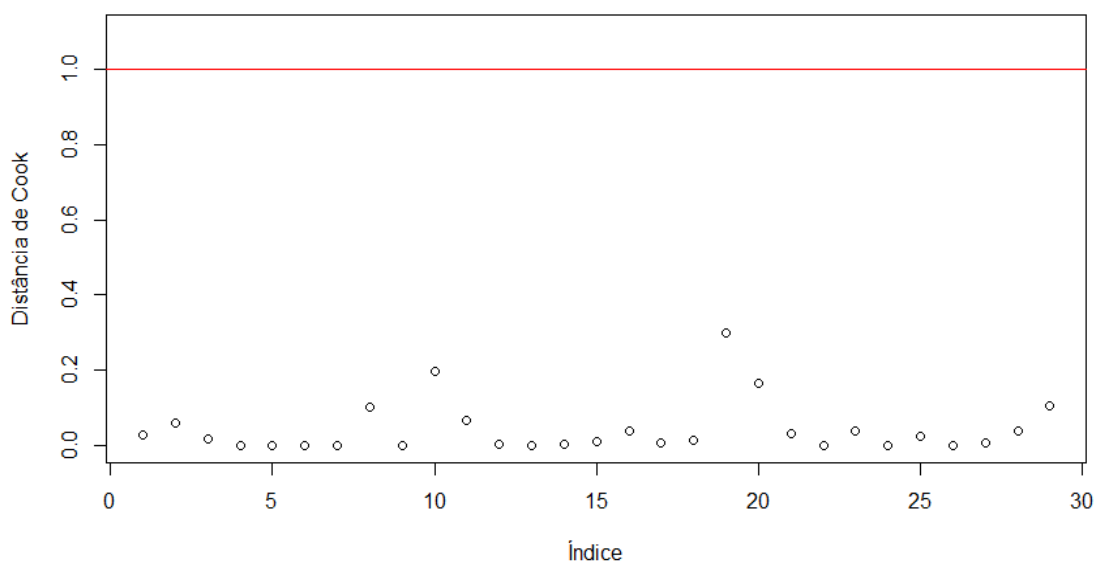


Figura 16: Gráfico QQ, resíduos e distância de Cook - modelo final (Tabela 10)

Conclui-se portanto que sexo, gravidade dos sintomas de TOC tipo sexual, e os scores de dor e qualidade de vida geral são conjuntamente bons preditores da redução da escala YBOCS de compulsão.

○ Marcadores

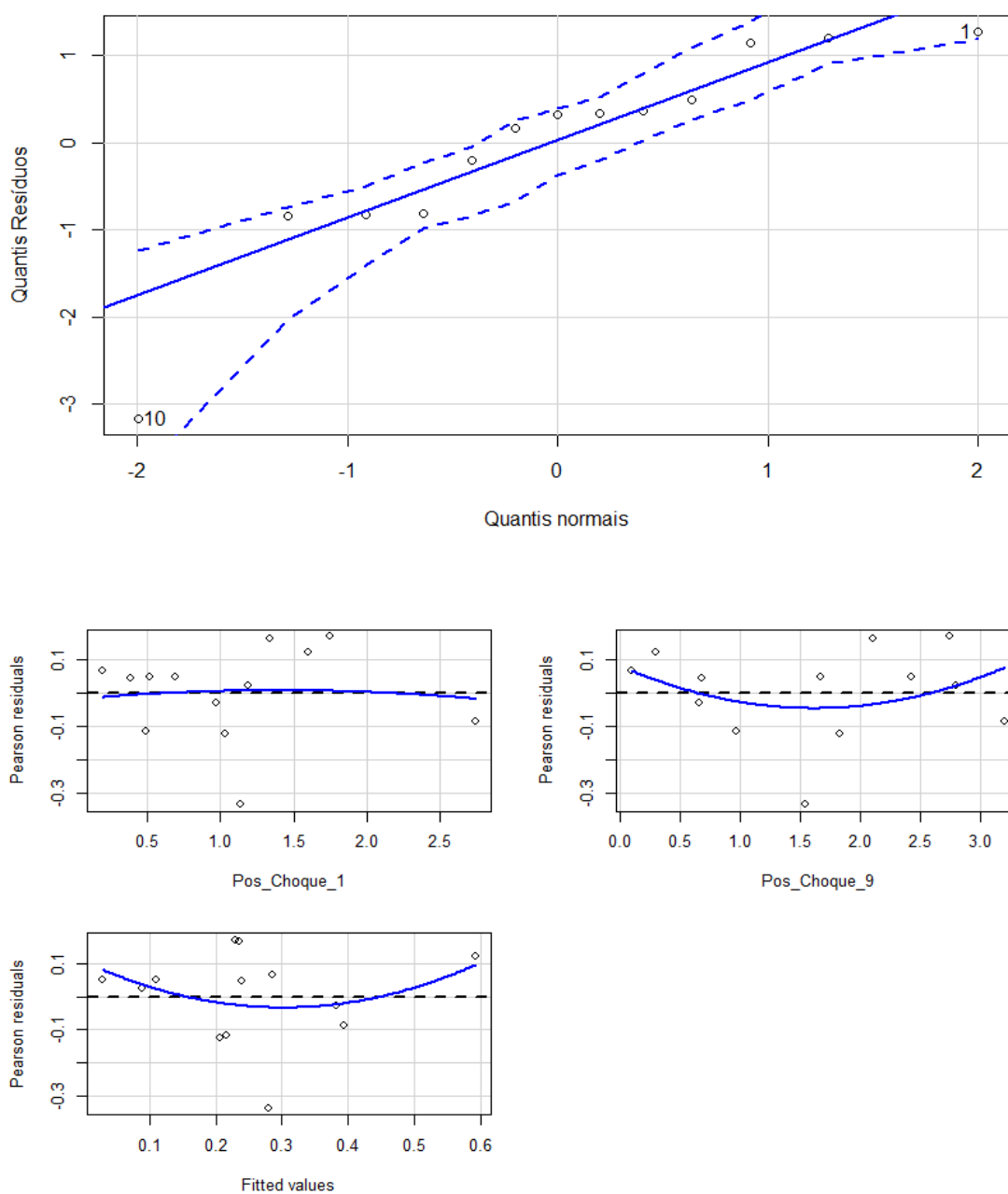
As tentativas de obtenção de um modelo para marcadores de redução da escala YBOCS de compulsão utilizando as variáveis selecionadas por *informative value* foram infrutíferas, mostrando-se necessária outra abordagem. Ajustou-se inicialmente um modelo com as 10 medidas de pós-choque, e a partir dele foi feita manualmente a seleção de variáveis de modo iterativo, resultando no modelo final abaixo, com apenas duas variáveis explicativas e ajuste satisfatório (Tabela 11):

Tabela 11: Modelo final de regressão linear para marcadores de compulsão

Variável	Estimativa	Erro Padrão	t value	Valor p
(Intercept)	0,257	0,088	2,907	0,016

Pos_Choque_1	0,240	0,081	2,982	0,014
Pos_Choque_9	-0,163	0,055	-2,986	0,014

A análise diagnóstica do modelo final consta na Figura 17.



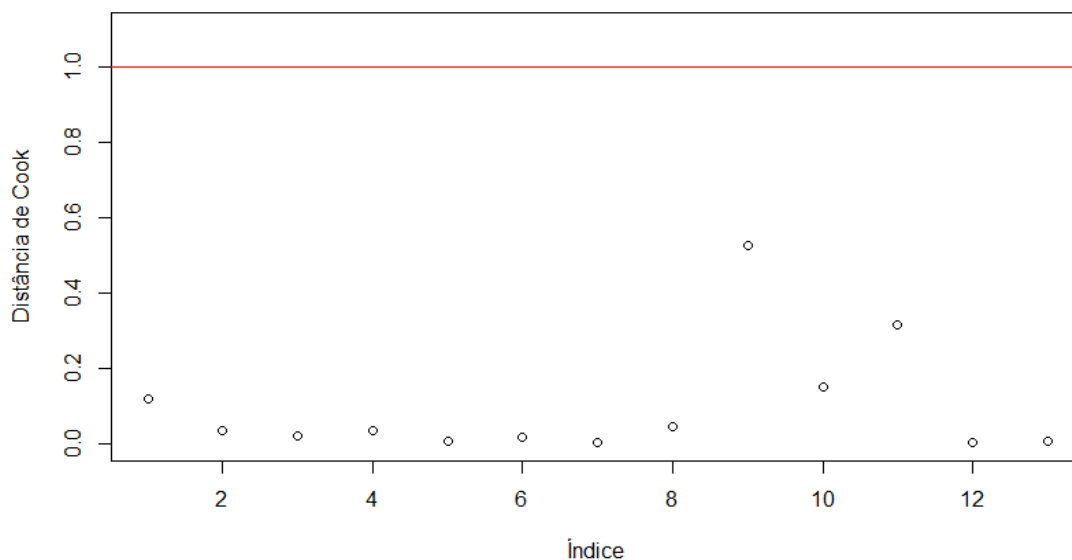


Figura 17: Gráfico QQ, resíduos e distâncias de Cook - modelo final (Tabela 11)

Conclui-se portanto que as medidas de pós-choque 1 e 9 são conjuntamente boas marcadoras de redução da escala YBOCS para compulsão.

6.2. Variável Categórica

Para a abordagem do problema considerando a variável resposta como categórica, foi utilizado o mesmo método de categorização do IV, ou seja, indivíduos que obtiveram uma redução na escala YBOCS superior a 20%, foram classificados como pacientes com melhora e os que não obtiveram essa redução como pacientes sem melhora.

Para a análise da variável de melhora foram utilizados modelos de regressão logística e a análise discriminante linear de Fisher (ADL). Estes modelos serão melhor discutidos nesta seção, inicialmente pela regressão logística e em seguida pela ADL.

• 6.2.1. Regressão logística para Obsessão

- Preditores

Como citado anteriormente, a pedido dos pesquisadores as variáveis sexo e idade foram consideradas nas etapas iniciais dos ajustes de preditores. No entanto, essas variáveis não se mostraram relevantes para o modelo de regressão logística.

Inicialmente, a seleção de variáveis foi feita de forma automática através do critério de informação de Akaike. Porém este método não encontrou variáveis relevantes e devido a isso, a seleção final foi feita de forma manual, através do valor p do teste de Wald, o que resultou no modelo descrito na Tabela 11.

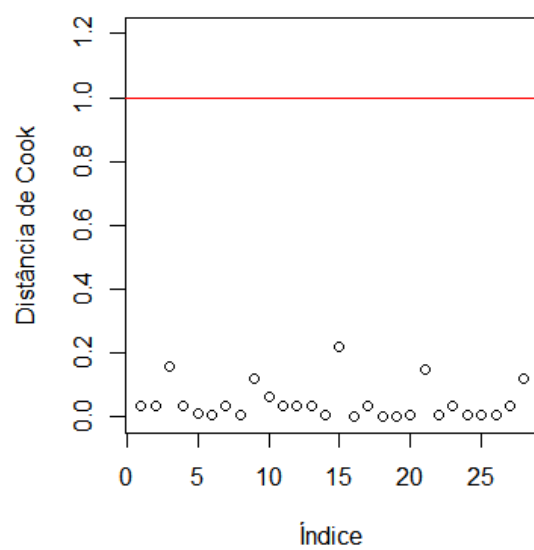
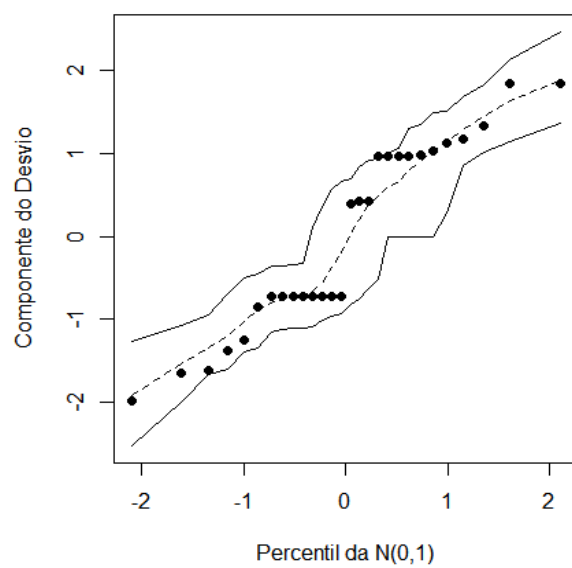
Tabela 11: Modelo final de regressão logística para preditores de obsessão

Variável	beta	exp(beta)	5%	95%	Valor p
(Intercept)	0,750	2,117	0,559	10,449	0,381
Grav_sexual	0,134	1,143	1,016	1,310	0,077
TOC_contaminacao (presente)	-2,065	0,127	0,019	0,604	0,043

Para os modelos de regressão logística, os parâmetros estimados (betas) dizem respeito ao logaritmo da chance de um evento em relação a outro, neste caso, a chance de não haver melhora em relação a haver melhora. A interpretação desses parâmetros é menos intuitiva, por conta disso, para os modelos de regressão logística será fornecida a interpretação da exponencial dos betas que fornece informações sobre a razão de chances do parâmetro.

Dessa forma, o aumento de 1 ponto na gravidade do TOC do tipo sexual aumenta em 14% a chance de não haver melhora em relação a chance de haver melhora do paciente. Além disso, a presença de TOC do tipo contaminação reduz a mesma chance em aproximadamente 88% vezes.

Com relação a qualidade do ajuste, observa-se através da Figura 18, que os desvios do modelo estão dentro do envelope simulado e ele não apresenta pontos influentes, segundo a distância de cook ($Di < 1$) (COOK e WEISBERG, 1982). Ademais, a área sob a curva ROC é de 0.776 o que indica um bom ajuste segundo (Hosmer et al., 2013).



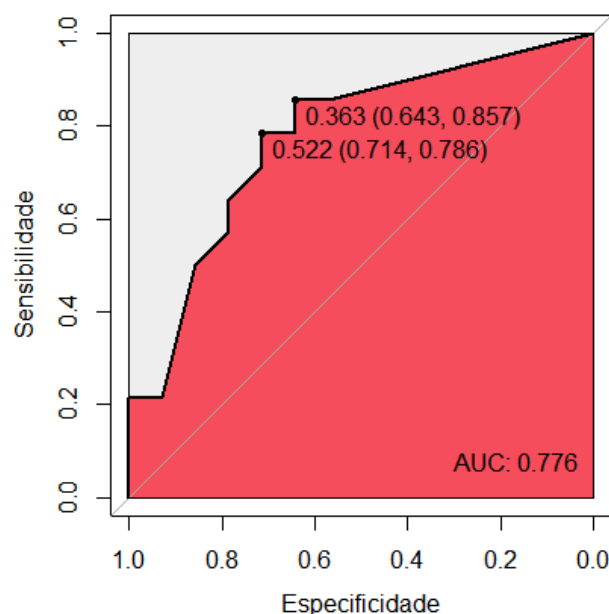


Figura 18: Envelope simulado, distância de cook e curva ROC - modelo final (Tabela 11)

- Marcadores

As variáveis selecionadas inicialmente pelo IV, não se mostraram relevantes nos métodos de seleção de variável, seja pelo critério de Akaike ou pelo valor p de Wald, também foram testadas medidas resumo das variáveis de pós choque (mínimo, média e máximo), que não passaram pelos critérios de seleção de variáveis. Além disso, foram testadas as próprias medições de pós choque, que resultaram no modelo descrito na Tabela 12.

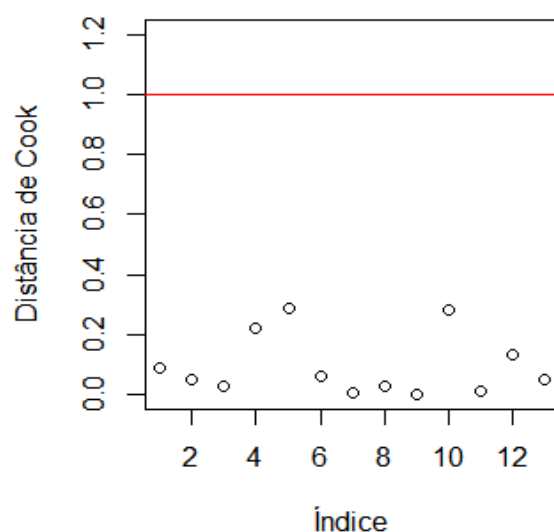
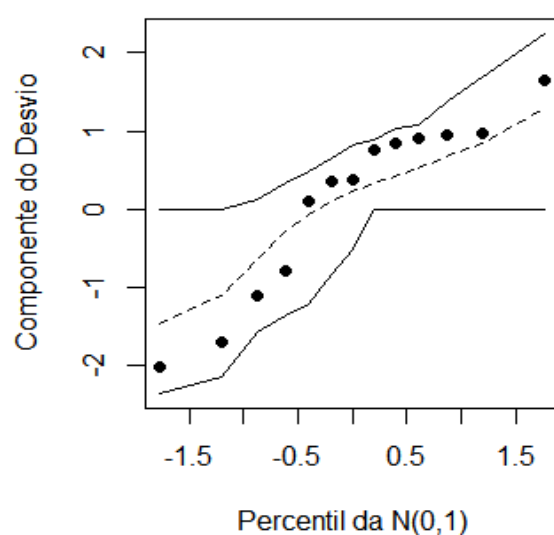
Tabela 12: Modelo final de regressão logística para marcadores de obsessão

Variável	beta	exp(beta)	5%	95%	Valor p
(Intercept)	1,209	3,350	0,282	80,463	0,455
Pos_Choque_1	2,963	19,360	1,227	> 1.000	0,163
Pos_Choque_9	-1,957	0,141	0,009	0,866	0,138

Apesar dos parâmetros apresentarem valores p relativamente altos, eles estão abaixo do nível de 10% do valor p adaptado. Uma interpretação possível do modelo é

que a chance de não haver melhora em relação a haver melhora é 19 vezes maior para o paciente que tenha 1 ponto a mais de condutância de pele para o choque 1, fixando-se a reação ao choque 9. Ao fixar o pós choque 1 temos que, a mesma chance é 7 vezes menor para o paciente que tenha um ponto a mais na condutância de pele do nono choque.

De forma análoga ao modelo com variáveis preditoras, este modelo também possui um bom ajuste, como é possível observar na Figura 19.



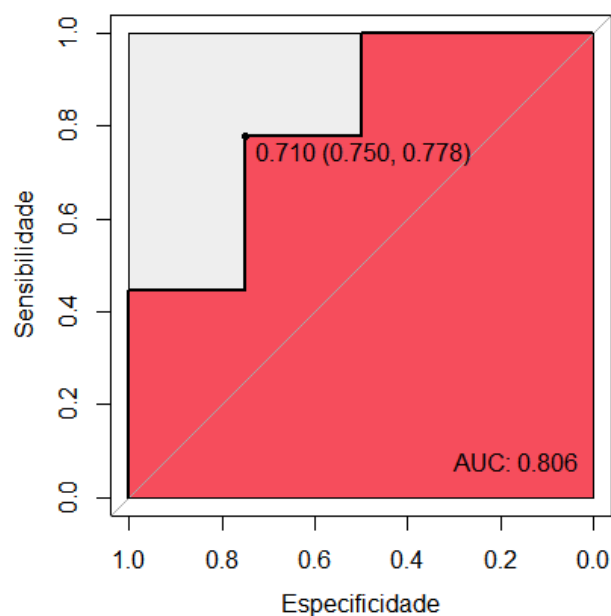


Figura 19: Envelope simulado, distância de cook e curva ROC - modelo final (Tabela 12)

• 6.2.2. Regressão logística para Compulsão

○ Preditores

Para os preditores de compulsão, os procedimentos de seleção de variáveis foram análogos ao de obsessão, dentre eles o que obteve o melhor modelo foi o critério de informação de Akaike. O modelo resultante deste critério é descrito na Tabela 13.

Tabela 13: Modelo final de regressão logística para preditores de compulsão

Variável	Estimativa	exp(beta)	5%	95%	Valor p
(Intercept)	-0,234	0,791	0,244	2,489	0,735
ABIPEMEC	-1,777	0,169	0,033	0,699	0,051
Grav_sexual	0,121	1,129	1,005	1,290	0,102

Para este modelo, o paciente pertencente a classe social C, tem a chance de não haver melhora em relação a haver melhora, aproximadamente 6 vezes menor do que um paciente de outra classe social. Além disso, o aumento de 1 ponto na gravidade do TOC do tipo sexual aumenta em 13% a mesma chance.

Da mesma forma que os modelos para obsessão anteriores, este modelo também apresenta uma boa qualidade de ajuste, como pode ser visto na Figura 20.

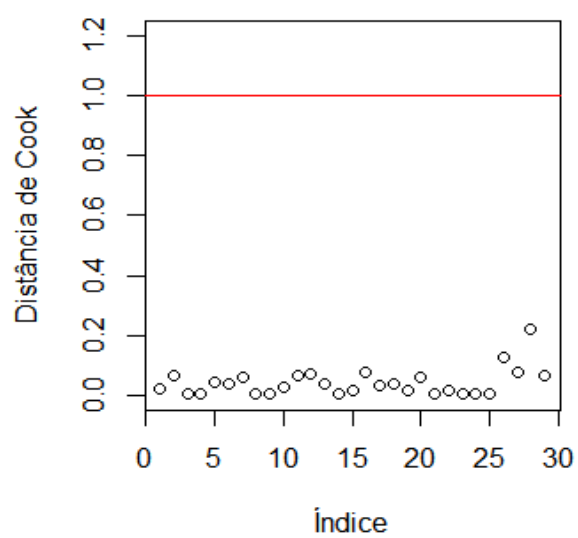
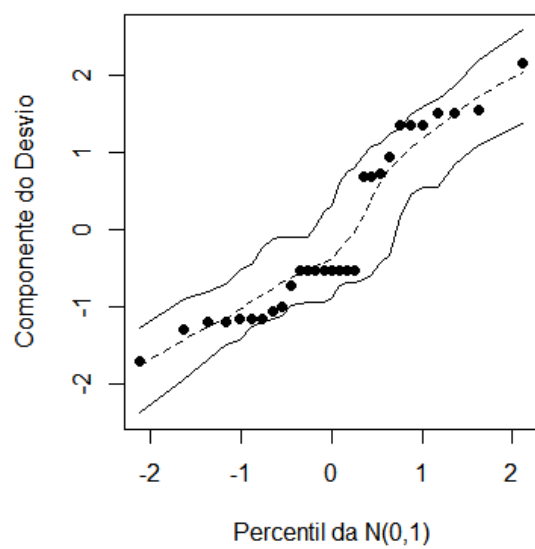
- Marcadores

Para os marcadores de compulsão, não foi possível encontrar uma combinação de variáveis que passasse pelos critérios de seleção de variáveis. Em todos os modelos o valor p de Wald para as variáveis se mostrou razoavelmente alto e, além disso, os modelos resultantes não apresentavam bons ajustes aos dados.

A Tabela 14 descreve o modelo que melhor se ajustou aos dados e a Figura 21 mostra a distância de Cook deste modelo, antes e após a retirada de uma observação influente.

Tabela 14: Modelo final de regressão logística para marcadores de compulsão

Variável	Estimativa	exp(beta)	5%	95%	Valor p
(Intercept)	-1,153	0,316	1,087	3,777	0,479
Pos_Choque_1	14,992	> 1.000	59,941	> 1.000	0,120
Pos_Choque_2	-15,698	< 0,001	< 0,001	0,030	0,130



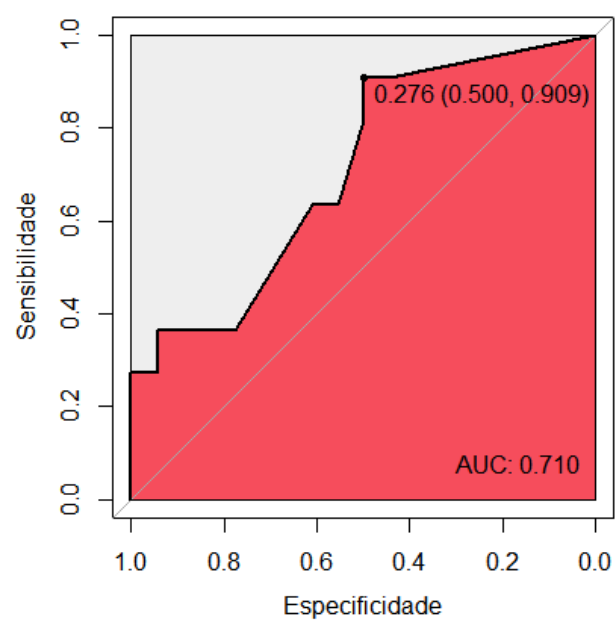
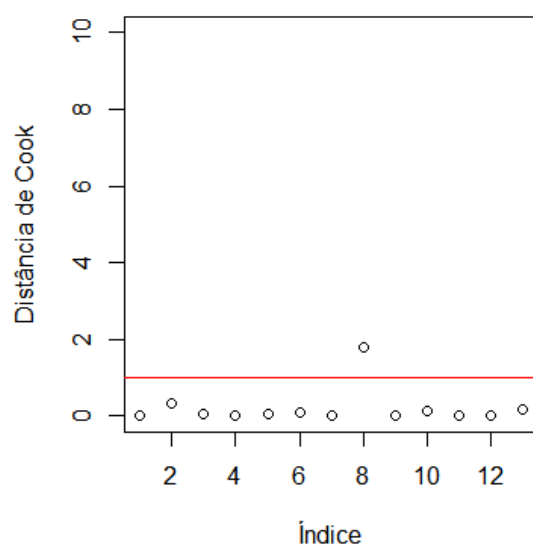


Figura 20: Envelope simulado, distância de cook e curva ROC - modelo final (Tabela 13)



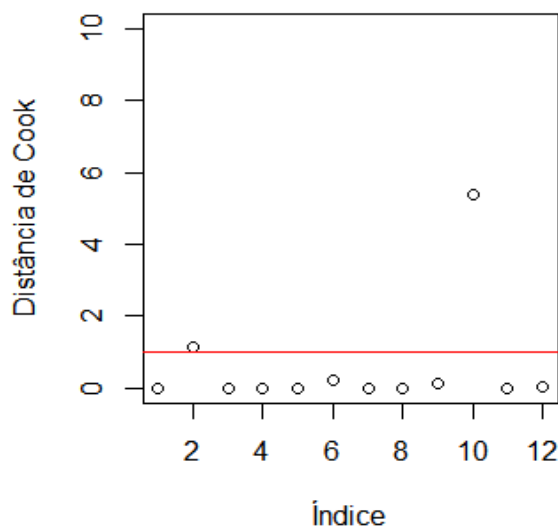


Figura 21: Distância de cook antes e após remoção de ponto influente - modelo final (Tabela 14)

• 6.2.3. Análise Discriminante para obsessão e compulsão

Para a realização da análise discriminante foi definido um critério de seleção de variáveis, descrito nas etapas a seguir:

1. Padroniza-se os dados entre 0 e 1.
2. Inicia-se o modelo com todas as variáveis sugeridas.
3. Observa-se a matriz de confusão, sensibilidade e especificidade do modelo, e gráficos de densidade estimada dos valores do discriminante.
4. Retira-se a variável com menor coeficiente da combinação.
5. Caso o item 3 não apresente piora, repete-se o item 4. Caso contrário, a variável é reincorporada, e obtemos o modelo final.

De tal modo, serão apresentados os modelos finais, juntamente com as matrizes de confusão, curvas ROC, e as densidades estimadas.

- Obsessão
 - Preditores

O modelo final está apresentado na Tabela 15.

Tabela 15: Modelo final de análise discriminantes para preditores de obsessão

Variável	LD1
Escolaridade2	-1,532
Escolaridade5	-3,635
Escolaridade6	-1,735
Escolaridade7	-4,230
Escolaridade8	-4,181
TOC_contaminacao1	-2,059
min_choque	1,924
Grav_ansiedade	-2,387
TOC_acumulacao1	1,566

Sua matriz de confusão e medidas de desempenho são apresentadas na Tabela 16 e 17.

Tabela 16: Matriz de confusão - modelo final (Tabela 15)

	Melhora	Não	Sim
Não		11	0
Sim		3	14

Tabela 17: Medidas de desempenho - modelo final (Tabela 15)

Sensitivity	0,786
Specificity	1
Balanced Accuracy	0,893

Utilizando a divisão padrão do modelo em questão, classificou-se corretamente 100% das melhoras (especificidade), e 78,6% dos casos em que não houve melhora.

A curva ROC gerada para o modelo é apresentada na Figura 22, o valor próximo de 1 da área abaixo da curva ROC, alinhado à matriz de confusão e à acurácia, contribui para demonstrar o bom ajuste do modelo. A linha vertical vermelha no gráfico de densidades representa o valor destacado na curva ROC.

Conclui-se então que a escolaridade, presença dos TOC's tipo contaminação e acumulação, o valor mínimo das variáveis choque, e a gravidade dos sintomas de ansiedade são bons preditores de melhora para obsessão, segundo o modelo.

Os valores apresentados na Tabela 15 são os pesos dados a cada uma das variáveis na função discriminante (que classifica como melhora ou não melhora). Assim sendo, é possível estabelecer variáveis que possuem influência mais forte na melhora dos pacientes, de acordo com o valor absoluto do coeficiente. As análises discriminantes apresentadas a seguir podem ser interpretadas de modo análogo.

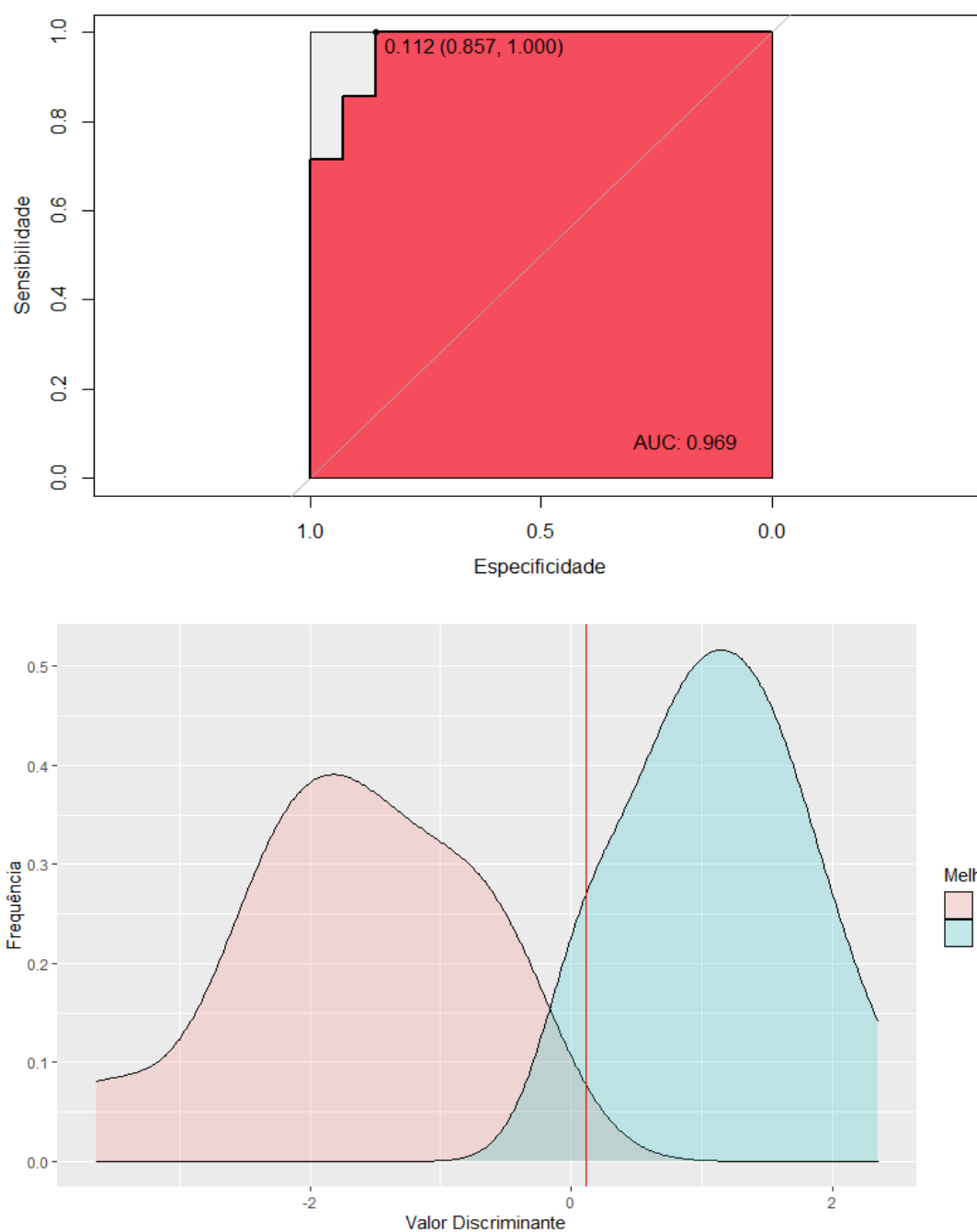


Figura 22: Curva ROC e densidade estimada - modelo final (Tabela 15)

○ Marcadores

Para os marcadores, o tamanho amostral cai para 13 observações, e com ele vai-se parte da acurácia. O modelo final consta na tabela 18.

Tabela 18: Modelo final de análise discriminantes para marcadores de obsessão

Variável	LD1
Renewal_Mean_diff_CS_U	-3,652
Recall_Mean_diff_CS_U	2,338

Tabela 19: Matriz de confusão - modelo final (Tabela 18)

	Melhora	Não	Sim
Não		3	2
Sim		5	8

Tabela 20: Medidas de desempenho - modelo final (Tabela 18)

Sensitivity	0,375
Specificity	0,800
Balanced Accuracy	0,588

O modelo possui uma acurácia pior, mas ainda se demonstra eficiente para classificar corretamente os casos em que houve melhora (especificidade). A curva ROC e as densidades também acusam a menor acurácia (Figura 23).

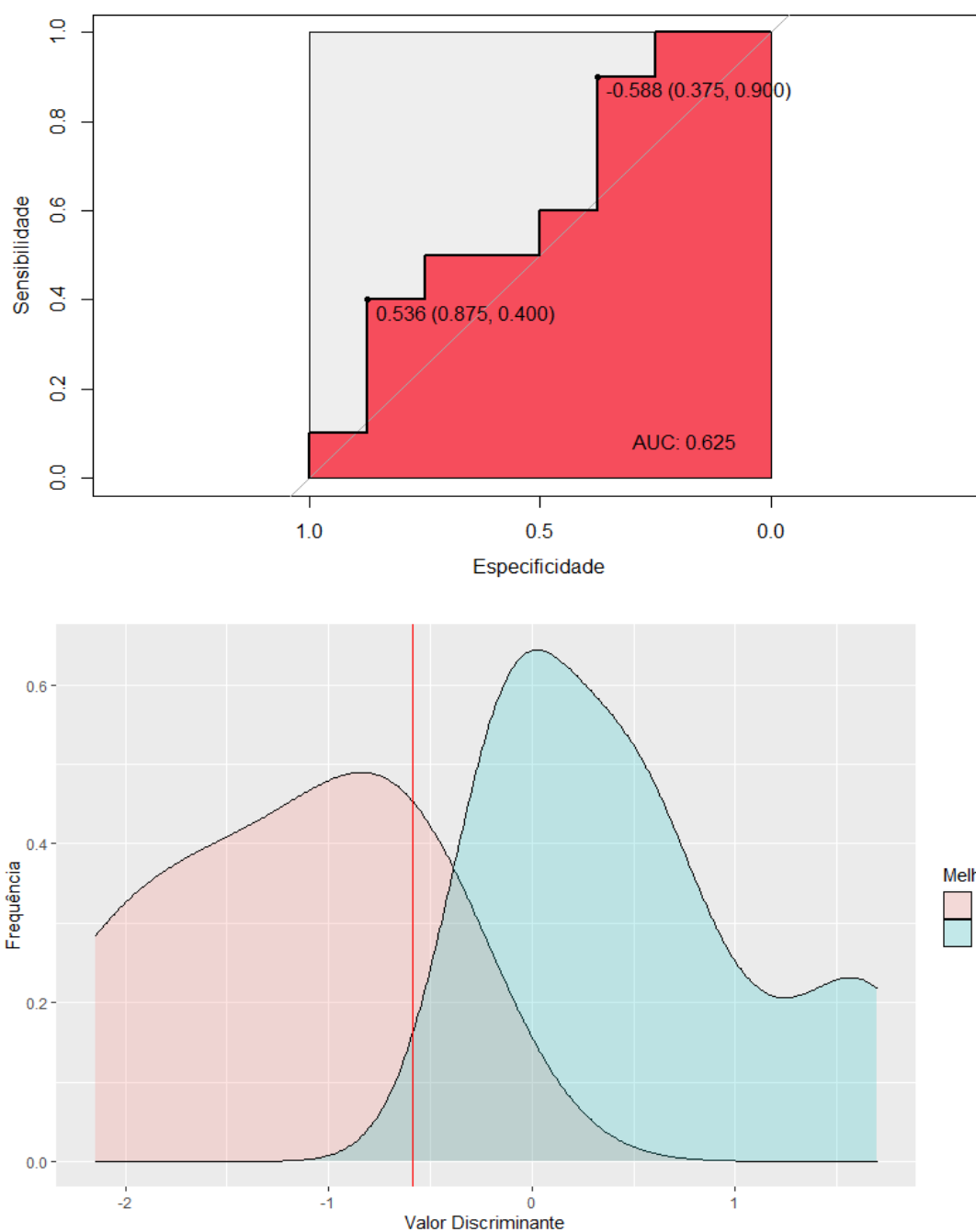


Figura 23: Curva ROC e densidades estimadas - modelo final (Tabela 18)

Conclui-se portanto que as variáveis de renovação e recuperação na ausência de choque são marcadoras de melhora para obsessão.

- Compulsão
 - Preditores

De modo análogo, segue o modelo final para preditores de compulsão, na Tabela 21.

Tabela 21: Modelo final de análise discriminantes para preditores de compulsão

Variável	LD1
Sexo2	-1,227
ABIPEME3	-1,220
ABIPEME4	-0,156
Grav_ansiedade	-0,417
Grav_sexual	1,782
QV_geral	2,982
QV_lim_fisica	-2,300

Tabela 22: Matriz de confusão - modelo final (Tabela 21)

	Melhora	Não	Sim
Não		18	4
Sim		0	7

Tabela 23: Medidas de desempenho - modelo final (Tabela 21)

Sensitivity	1
Specificity	0,636
Balanced Accuracy	0,818

Embora haja uma disparidade maior entre sensibilidade e especificidade, a acurácia ainda é satisfatória, e o modelo consegue classificar com eficiência a melhorar, o que também é demonstrado nos gráficos da Figura 24.

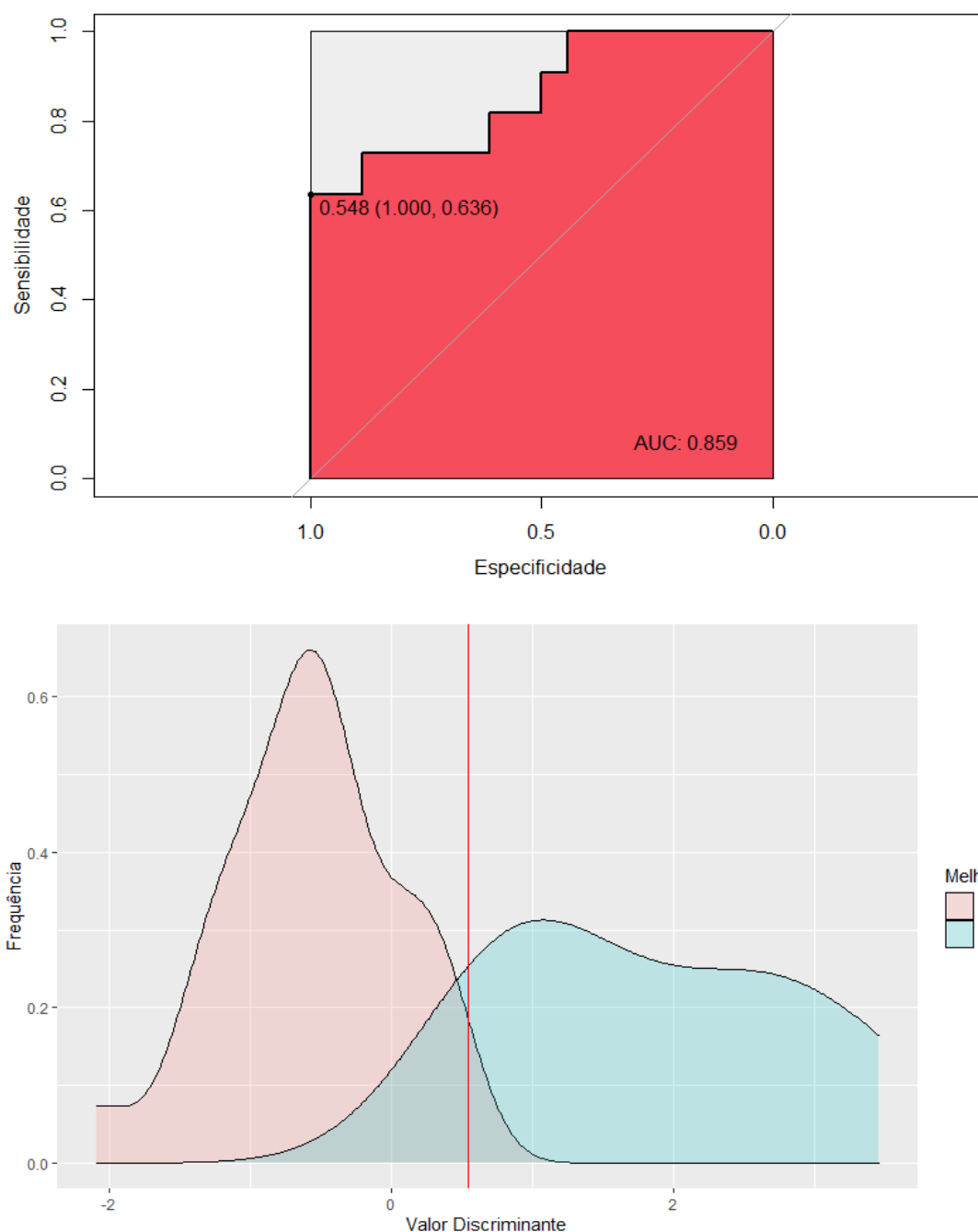


Figura 24: Curva ROC e densidade estimada - modelo final (Tabela 21)

Mais uma vez, a área abaixo da curva aliada à separação visualmente representada pelas densidades e às medidas de desempenho indicam bom ajuste.

Conclui-se então que sexo, classe social, gravidade dos sintomas de ansiedade e TOC do tipo sexual, e os escores de qualidade de vida geral e limitação física são bons preditores de melhora para compulsão.

- Marcadores

Analogamente a o que foi feito para a regressão logística e regressão linear, foram utilizadas as variáveis de pós-choque em detrimento às medidas resumo. O modelo final é dado na Tabela 24.

Tabela 24: Modelo final de análise discriminantes para marcadores de compulsão

Variável	LD1
Pos_Choque_1	10,884
Pos_Choque_3	-14,781
Pos_Choque_4	15,451
Pos_Choque_9	-7,840

Tabela 25: Matriz de confusão - modelo final (Tabela 24)

	Melhora	Não	Sim
Não	5	0	
Sim	1	7	

Tabela 26: Matriz de confusão - modelo final (Tabela 24)

Sensitivity	0,833
Specificity	1
Balanced Accuracy	0,917

O modelo apresentou boa acurácia, tendo classificado apenas uma observação de modo incorreto. A curva ROC e as densidades acumuladas também demonstram o bom desempenho.

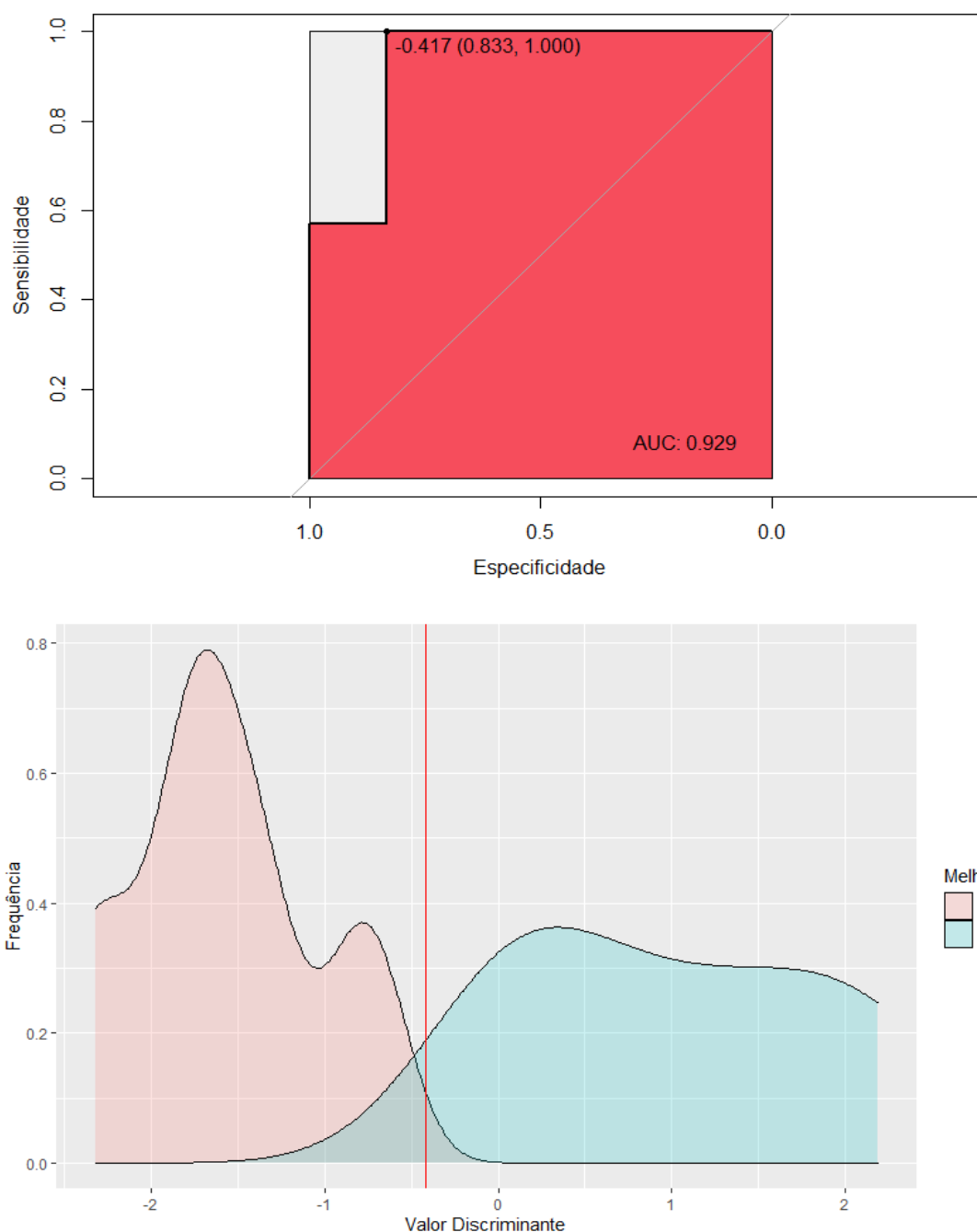


Figura 25: Curva ROC e densidades estimadas - modelo final (Tabela 24)

Concluimos então que as variáveis de pós-choque 1, 3, 4 e 9 são marcadores de melhora para compulsão.

7. Considerações Finais

O ajuste dos modelos identificou bons candidatos para marcadores e preditores, tanto da melhora dos pacientes, quanto da redução da escala YBOCS, para obsessão e compulsão. É importante deixar claro que, embora os modelos possuam bom ajuste, o tamanho amostral em muitos casos é pequeno, indicando que as variáveis apontadas pelo modelo não devem ser tomadas cegamente, mas sim como um norteador.

É importante também ter em mente que ainda há informações de neuroimagem a serem adicionadas ao banco de dados. A obtenção desses dados é bem complexa e, por conta disso, eles não foram disponibilizados a tempo de serem incluídos na análise, e assim que estiverem disponíveis poderão ser adicionados, incrementando a qualidade dos marcadores e preditores.

A Figura 26 apresenta um resumo das variáveis apontadas como boas preditoras e marcadoras, de acordo com o número de modelos que as considerou relevantes. Assim, o círculo verde ao lado da variável mostra que ela foi dada como significativa pelas três técnicas adotadas. O círculo amarelo indica seleção em duas técnicas, e por fim o círculo vermelho indica seleção em apenas uma técnica.

Dessa forma, as variáveis mais relevantes como preditores para os marcadores de melhora de obsessão foram a presença de TOC do tipo contaminação, gravidade de TOC do tipo sexual, gravidade da ansiedade e o mínimo das medidas de choque. Para os marcadores de melhora de compulsão as mais relevantes foram gravidade de TOC do tipo sexual, sexo, índice de qualidade de vida geral e classe social.

Para os marcadores da melhora de obsessão, a variável mais relevante foi a diferença entre os estímulos com e sem choque não extintos, e para compulsão as mais relevantes foram as medidas pós-choque 1 e pós-choque 9.

Preditores

Obsessão	Compulsão
TOC Contaminação	Gravidade TOC Sexual
Gravidade TOC Sexual	Sexo
Gravidade Ansiedade	QV - Geral
Mínimo Choque	ABIPEME
Gravidade TOC Simetria	QV - Dor
Gravidade Depressão	QV - Limitação Física
QV - Bem-estar	
QV - Dor	
QV - Saúde física	
Presença TOC Acumulação	

Marcadores

Obsessão	Compulsão
Renovação CS+U - CS-U	Pós-Choque 1*
Pós-Choque 1	Pós-Choque 9
Pós-Choque 9	Pós-Choque 3
Recuperação CS+U - CS-U	Pós-Choque 4
	Pós-Choque 2*

Figura 26: Resumo dos marcadores e preditores selecionados.