

Aplicação da Teoria da Resposta ao Item ao score NESSCA de avaliação da progressão da Doença de Machado Joseph

Tássia Henkes Maciel

Professora Msc. Vanessa Leotti Torman

Professora Dra. Suzi Alves Camey

Porto Alegre, 21 de Dezembro de 2012.

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

Aplicação da Teoria da Resposta ao Item ao escore NESSCA de avaliação da progressão da Doença de Machado Joseph

Tássia Henkes Maciel

Monografia apresentada para obtenção
do grau de Bacharel em Estatística.

Banca Examinadora:
Professora Msc. Vanessa Leotti Torman
Professora Dra. Suzi Alves Camey
Professor Dr. João Riboldi

Porto Alegre, 14 de Janeiro de 2013.

Dedico este trabalho àqueles (quase idealistas) que, com muito empenho, divulgam, publicam, discutem, participam e compartilham seus conhecimentos pelo simples prazer de compartilhar.

Wanting to be able to

*"Impossibilities" are good
not to attach that label to;
since, correctly understood,
if we wanted to, we would
be able to be able to*

Piet Hein

Agradecimentos

Aos meus pais, Lauro e Clélia, e à minha irmã, Pricila, pelo amor incondicional e por todas as vezes que entenderam minha ausência.

Aos Balotins da minha vida, por serem também meus pais e irmãos.

A todos os familiares, especialmente à vó Tereza, por essa família linda, e ao Dindo Nuno, por me querer tão bem.

Aos amigos e colegas do Hospital de Clínicas de Porto Alegre, por todo conhecimento que compartilharam comigo de forma tão gentil.

Aos amigos e colegas do Grupo RBS, que me ensinam a *ir além* todos os dias.

Aos colegas de curso, pelo incentivo sempre. De uma forma muito especial à Pricila, Sabrina e Natalia, por todos os domingos de sol em que estávamos estudando juntas.

Um abraço bem apertado para minhas orientadoras professoras Suzi e Vanessa, obrigada por tudo!

E o último agradecimento é para Josean, meu amor, por dividir comigo todos os momentos deste trabalho e da minha vida.

Resumo

A escala NESSCA (Neurological Examination Score for Spinocerebellar Ataxia) foi criada para medir o comprometimento de pacientes com ataxias spinocerebrais, entre elas a ataxia spinocerebral tipo 3, conhecida como Doença de Machado Joseph (DMJ). A escala é baseada em exames neurológicos e o escore pode estar entre 0 e 40. Nosso objetivo foi avaliar a NESSCA através da Teoria da Resposta ao Item (TRI). Este trabalho foi dividido em duas etapas: na primeira se avaliou o escore NESSCA obtido por diferentes modelos da TRI e em seguida foram aplicados os modelos calibrados na etapa anterior em outro conjunto de dados. Os resultados mostraram que os itens Câimbra e Vertigem não ajudam a medir o comprometimento com a DMJ e poderiam ser retirados da NESSCA. Outra melhoria proposta foi em relação ao agrupamento de categorias de resposta para alguns itens. Entre os modelos de resposta gradual de Samejima ajustados, um deles se destacou por ter altas correlações com a NESSCA (0,911, p -valor $<0,05$), alta consistência interna (Alfa de Cronbach 0,81), ser composto apenas por 12 itens e perder pouca informação em relação ao modelo completo (6,5%). A utilização da TRI em estudos futuros para avaliar o comprometimento com DMJ é recomendada, pois este método apresentou vantagens em relação ao método atualmente utilizado, especialmente em relação ao menor número de itens necessários para medir o traço latente e o ganho em discriminação ao utilizar uma escala contínua.

Sumário

Agradecimentos	5
Resumo.....	6
Sumário	7
Introdução.....	9
Metodologia.....	10
Pacientes e descrição da NESSCA.....	10
Análise estatística e modelo TRI escolhido	11
Resultados.....	11
Primeira etapa: calibração	12
Segunda etapa: validação	16
Discussão.....	17
Referências Bibliográficas	19
Anexos.....	21
Anexo 1: Escala NESSCA	21
Anexo 2: Gráficos do Modelo TRI Completo.....	23
Anexo 3: Estimativas pontuais dos parâmetros dos modelos TRI ajustados.....	29

Este artigo será submetido à revista European Journal of Neurology.

Introdução

O termo *análise de estrutura latente* foi criado por Paul F. Lazarsfeld para descrever uma classe de modelos matemáticos para estudos de comportamento, através de atributos utilizados como indicadores. Lazarsfeld desenvolveu muitos modelos durante a década de 1940, fundamentais para estabelecer este conceito teórico de uma variável não mensurável [1]. A partir da década de 50, muitos autores (como Lord, Samejima e Bock) contribuíram para a criação de novos modelos, com diferentes estruturas de respostas. Atualmente, termos como “traço latente” e “variável latente”, oriundos destes estudos, são utilizados para definir uma característica ou habilidade que não pode ser medida de forma direta [2]. Exemplos de variáveis latentes são: o nível de satisfação de um cliente, a habilidade de um aluno em matemática, o grau de depressão, a intensidade da dependência a uma droga, entre outros. Escalas adequadas são desenvolvidas para mensurar variáveis latentes e a análise das respostas pode ter duas abordagens distintas: a Teoria Clássica dos Testes (TCT) e a Teoria da Resposta ao Item (TRI).

A Teoria Clássica dos Testes é o modelo mais antigo (e ainda muito utilizado) para medir traços latentes e se baseia no escore do instrumento como um todo [3], frequentemente representado pela soma ou média dos itens da escala. Quando um instrumento é elaborado à luz da TCT os parâmetros dos itens dependem da amostra em que eles foram calculados [4]. Ou seja, tanto as informações sobre os itens, suas correlações, erros padrões quanto a validade e consistência interna do questionário dependem da particular amostra estudada. Devido a limitações como estas, apesar dos modelos da TCT terem sido muito importantes em pesquisas e descobertas por mais de 80 anos, atualmente eles têm sido substituídos pelos modelos da TRI [5].

A Teoria da Resposta ao Item fornece modelos matemáticos para os traços latentes, propondo formas de representar a relação entre a probabilidade de um indivíduo dar certa resposta a um item, seu traço latente e características dos itens [6]. Uma das grandes vantagens da TRI sobre a TCT é o fato de ela permitir a comparação de populações, pois sua principal característica é considerar como elementos os itens e não o escore final [3]. Outro benefício da TRI é o princípio da invariância: os parâmetros dos itens não dependem do traço latente do respondente e os parâmetros dos indivíduos não dependem dos itens apresentados [3].

A TRI pode ser aplicada em diversas áreas, não apenas em psicomетria e educação. Um estudo [8] de 2011 utilizou a TRI para explorar escalas que medem comprometimento cognitivo de pacientes com esquizofrenia e para analisar quantos itens são necessários para medir déficits cognitivos. Os resultados apontaram, por exemplo, que apenas 10 dos 12 itens de um instrumento eram necessários para medir o traço latente, sugerindo que o instrumento poderia ser reduzido. A escala UHDRS, que mede a gravidade da doença de Huntington, foi avaliada [9] com TRI, e os dados apontaram que alguns itens parecem bons para medir diferentes intensidades da doença, mas outros itens apresentaram características problemáticas. Outro estudo [10] recente utilizou a TRI na escala ADAS-cog, que mede a disfunção cognitiva em pacientes com Alzheimer. Entre os resultados, os autores observaram que muitos pacientes tinham o mesmo escore total e visível diferença nas disfunções cognitivas, o que sugeriu que os escores obtidos pela TRI poderiam medir melhor o traço latente.

O objetivo deste trabalho é avaliar através da TRI um instrumento utilizado para medir o comprometimento de pacientes com ataxias. Ataxia é a condição de falta de coordenação dos movimentos, sinal comum para diversas categorias de doenças genéticas [11]. Entre estas doenças está a ataxia spinocerebral tipo 3 (SCA3), chamada de Doença de Machado Joseph (DMJ), causada por alterações no gene localizado no cromossomo 14q32.1 [11]. Esta ataxia é também conhecida como Doença Açoriana, pois os primeiros casos foram registrados na década de 1970 em famílias estadunidenses de origem luso-açoriana [12]. O quadro clínico possui diferentes subtipos, com sintomas característicos das ataxias: falta de coordenação motora, alterações oculares, dificuldades na fala. A DMJ é neurodegenerativa, progressiva e geralmente se manifesta na idade adulta, causando dependência devido ao quadro clínico, sem nunca alterar o intelecto [11,12]. Para medir o comprometimento do paciente com ataxias spinocerebrais, foi criada a escala NESSCA (*Neurological Examination Score for Spinocerebellar Ataxia*) [13], baseada em uma avaliação quantitativa de exames neurológicos padronizados e seu foco está as principais características das ataxias.

Este trabalho está dividido em duas etapas. Na primeira, pretende-se avaliar o escore NESSCA obtido por diferentes modelos da TRI, compará-los com o escore obtido pela abordagem clássica e propor melhorias para a escala. Na segunda etapa, o objetivo é realizar a validação dos escores obtidos através da aplicação os modelos calibrados na etapa anterior em outro conjunto de dados, comparando os escores obtidos com escores de outras escalas utilizadas para medir o comprometimento com a DMJ.

Metodologia

Pacientes e descrição da NESSCA

Os dados da primeira etapa deste trabalho foram obtidos em um estudo do Hospital de Clínicas de Porto Alegre (HCPA) envolvendo 106 pacientes diagnosticados com DMJ, acompanhados entre maio de 1995 e junho de 2005 [13]. Os pacientes entraram no estudo com diferentes níveis de comprometimento devido à DMJ e optou-se por selecionar aleatoriamente qual observação de cada indivíduo faria parte da amostra. Na segunda etapa, foram utilizados os dados basais de um ensaio clínico sobre a segurança e a eficácia do uso de carbonatos de lítio em 60 pacientes com DMJ [14].

Algumas escalas são utilizadas, como SARA, Barthel, BDI, SCAFI, CCFS e a NESSCA para medir alguns aspectos do comprometimento com ataxias spinocerebrais. A SARA (*Scale for the Assessment and Rating of Ataxia*), por exemplo, é escala direcionada à severidade física da DMJ, com componentes axiais, membros, visão e fala, BDI (*Beck Depression Inventory*) está relacionado com sintomas de depressão e a escala de Barthel é utilizada para medir habilidades funcionais do dia-a-dia [13]. Os escores dos modelos TRI serão correlacionados com os escores destas escalas.

A escala NESSCA é composta por dezoito itens, onde treze correspondem a exames neurológicos e cinco são informados pelo paciente ou responsável, resultando em um escore total (TCT) entre zero e quarenta. A tabela do Anexo 1 apresenta a escala NESSCA completa. É possível observar que as respostas são ordinais e o número de categorias não é o mesmo para todos os itens.

Análise estatística e modelo TRI escolhido

De uma forma geral, os modelos da TRI dependem de três fatores: a natureza do item (dicotômicos ou politômicos), o número de populações envolvidas e da quantidade de traços latentes que está sendo medida [2,3]. No caso da escala NESSCA, os itens são politômicos e ordinais. Entende-se que neste contexto existe uma população envolvida e apenas um traço latente está sendo medido. Com estas características, o modelo TRI escolhido foi o modelo de resposta gradual de Samejima [16], que considera as respostas dos itens ordenadas entre si e a probabilidade de um indivíduo j escolher uma particular categoria k ou outra mais baixa do item i pode ser dada por

$$\text{logit}(\gamma_{ijk}) = \ln\left(\frac{\gamma_{ijk}}{1 - \gamma_{ijk}}\right) = \beta_i(\theta_j - \beta_{ik})$$

onde $i = 1, 2, \dots, I$ o número de itens; $k = 0, 1, 2, \dots, m_i$ número de categorias de cada item i ; $j = 1, 2, \dots, n$, sendo n o número de sujeitos pesquisados; γ_{ijk} é a probabilidade acumulada do paciente j que possui θ_j comprometimento com a doença, responder o item i com categoria k ou menor; β_i é o parâmetro de discriminação do i -ésimo item; θ_j representa o comprometimento do j -ésimo paciente com a DMJ (traço latente); β_{ik} é o parâmetro de dificuldade da k -ésima categoria do item i .

Neste trabalho, o escore do comprometimento com a DMJ obtido pela soma dos itens da escala NESSCA será chamado apenas de NESSCA. Os modelos da TRI apresentados são o Modelo Completo e os modelos 1 a 6. O Modelo Completo é composto por todos os 18 itens da escala NESSCA e a partir dele será ajustado o Modelo 1, com algumas modificações. Os Modelos 2 a 6 são obtidos por modificações do Modelo 1, retirando-se sequencialmente itens com menores discriminações. Os escores destes modelos serão chamados de escores TRI, e representam as estimativas da TRI para o comprometimento com a DMJ. Para todos os modelos ajustados, as estimativas pontuais dos parâmetros dos itens estão apresentadas no Anexo 3.

Para comparar os escores e modelos foram utilizados Alfa de Cronbach, correlação de Pearson, gráficos de dispersão e de Bland-Altman. As análises foram feitas no software R versão 2.15.1 [17] com apoio do pacote ltm [18].

Resultados

Cada etapa das análises foi realizada em um conjunto de dados diferente. Para a etapa de análise dos itens da NESSCA e calibração dos itens, a amostra era composta por 106 pacientes, 54 (50,9%) do sexo masculino e 52 (49,1%) do sexo feminino. A idade média do grupo foi 43,6 anos (DP=12,3 anos). Para a segunda etapa das análises, a amostra era composta por 60 pacientes, a idade média do grupo foi 40,9 anos (DP=9,1 anos). Quanto ao gênero, 29 (48,3%) pacientes eram do sexo masculino e 31(51,7%) do sexo feminino.

Primeira etapa: calibração

O primeiro modelo ajustado na amostra de 106 pacientes, que foi chamado de Modelo Completo, utilizou todos os 18 itens com suas categorias de resposta originais, exceto Ataxia de marcha (CERAX), que não teve nenhum respondente para a categoria 0 e por isso se considerou que as categorias 1,2,3 e 4 seriam, respectivamente, 0,1,2 e 3. Para este modelo, o Alfa de Cronbach foi 0,783, a informação total do teste foi 48,42 e a correlação entre os escores TRI e a NESSCA foi 0,960. As curvas características dos itens e a informação total do teste estão disponíveis no Anexo 2.

A primeira decisão tomada para melhorar a escala está relacionada ao agrupamento de categorias. As curvas características dos itens Nistagmo (NISTA), Fasciculações (FASC) e Perda sensorial (SENS) estão apresentadas nas Figuras 1 a 3. O eixo horizontal representa os comprometimentos com a doença considerando distribuição normal com média zero e desvio padrão igual a 1. Para NISTA, podemos observar que a probabilidade de responder a categoria 3 é muito pequena em relação às outras categorias, e ocorre apenas para comprometimentos acima de três desvios da média, o que sugere que a categoria 3 pode ser agrupada com a 2. Para os itens FASC e SENS, a probabilidade de responder a categoria 3 é sempre menor que a probabilidade de responder outra categoria, o que também sugere que as categorias 2 e 3 podem ser agrupadas.

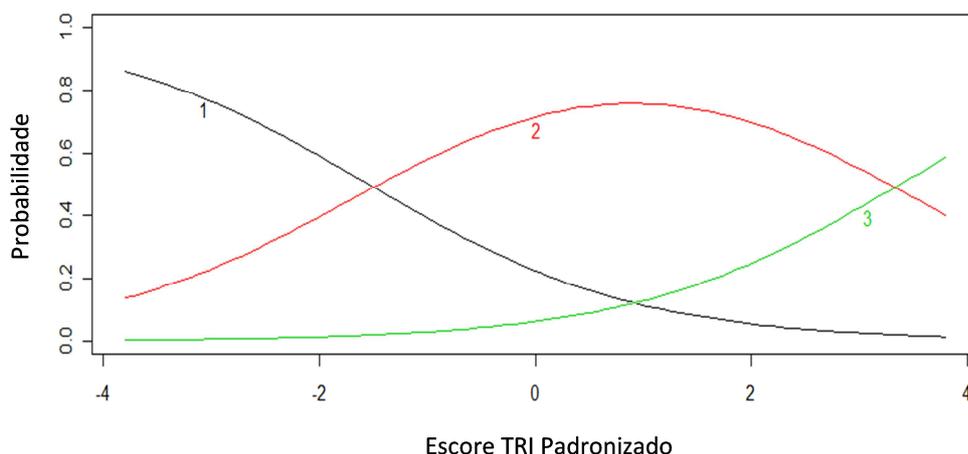


Figura 1: Curva característica do Item NISTA Modelo Completo

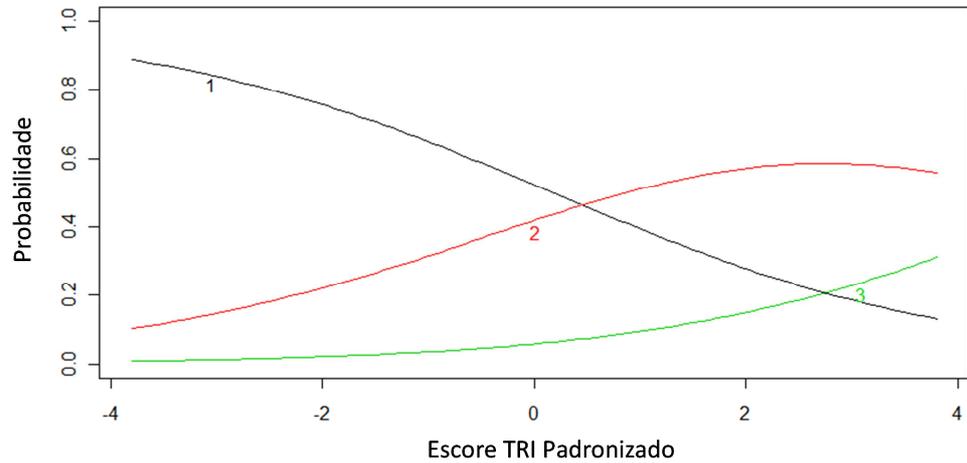


Figura 2: Curva Característica do Item FASC Modelo Completo

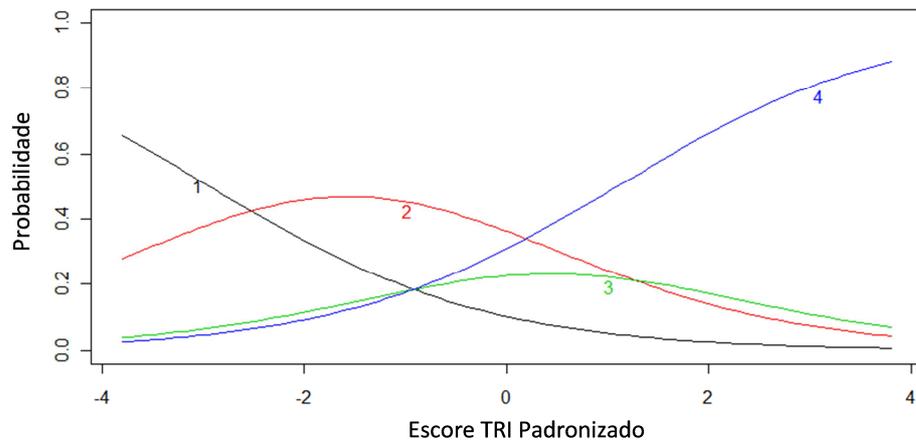


Figura 3: Curva Característica do Item SENS Modelo Completo

Além destes agrupamentos, dois itens apresentaram problemas com relação à discriminação: Câimbra (CAIM) e Vertigem (VERT). O primeiro apresentou discriminação negativa, que viola o pressuposto da TRI de que quanto maior o traço latente, maior a probabilidade de responder categorias com pior gravidade, e o segundo teve discriminação igual a 0,024, considerada muito baixa. Portanto, estes itens foram excluídos.

Um novo modelo, que será chamado de Modelo 1, foi ajustado a partir destas decisões, e o escore da TRI foi transformado para ter a mesma média e desvio padrão da NESSCA (média 14,27 e desvio padrão 4,72). Os gráficos de dispersão e de Bland-Altman dos escores estão representados nas figuras 4 e 5. O gráfico de Bland-Altman possui no eixo horizontal a média entre escores NESSCA e TRI e no eixo vertical a diferença entre eles, e permite observarmos que as diferenças entre os escores chega a 4 pontos, sem existir padrão no comportamento delas (como um funil, por exemplo) que indicasse que estas diferenças são maiores ou menores dependendo do escore médio (eixo horizontal). A correlação de Pearson entre a NESSCA e o escores do Modelo 1 foi 0,962. A informação total dos itens no Modelo 1 foi 46,16 e o Alfa de Cronbach foi 0,814. O gráfico da informação total está na Figura 6, considerando escore TRI padronizado.

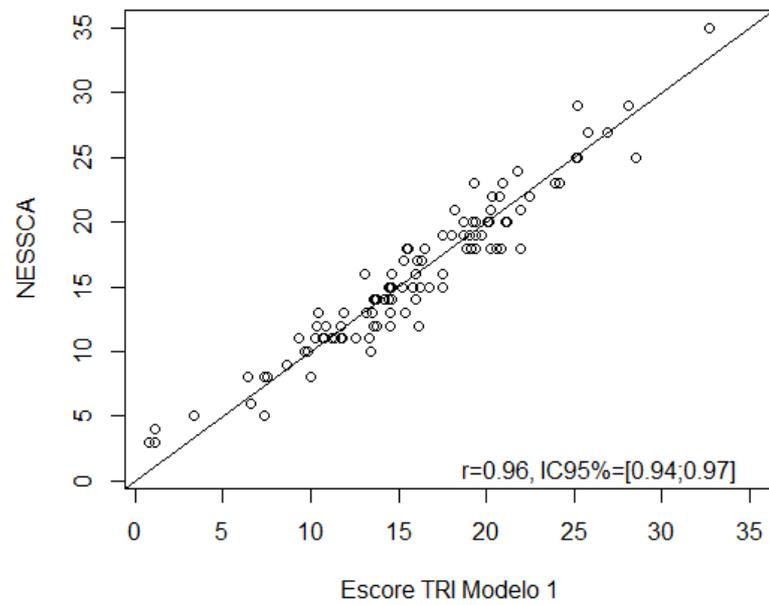


Figura 4: Diagrama de dispersão entre escores TRI e NESSCA

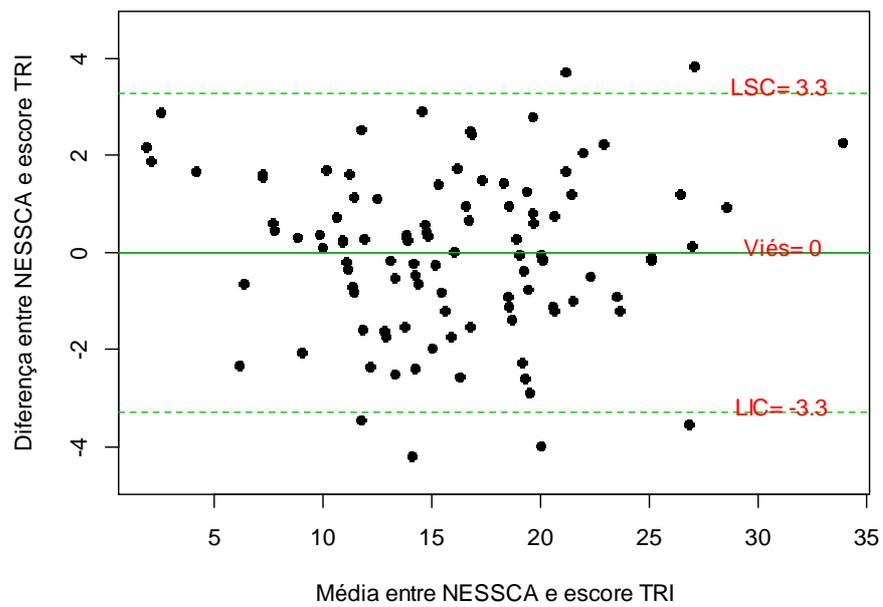


Figura 5: Gráfico de Bland-Altman para NESSCA e escore TRI Modelo 1

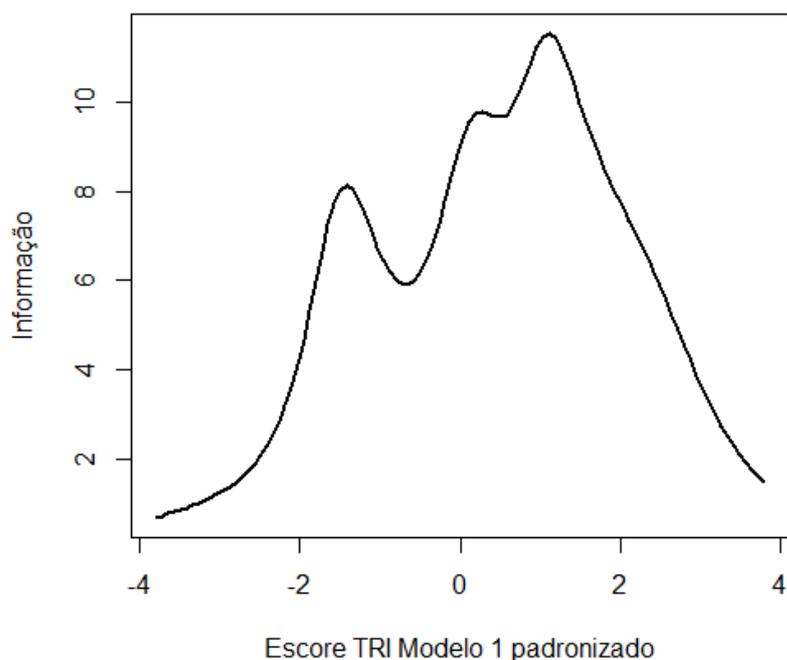


Figura 6: Informação total do Teste Modelo 1

A partir do Modelo 1, foram propostos os modelos 2 a 6, retirando sequencialmente o item com menor discriminação no modelo anterior, dentre os itens com discriminação menor que 1. Os resultados estão na Tabela 1. Do Modelo 1 para o Modelo 2, por exemplo, retirou-se o item Nistagmo (NISTA), pois possuía a menor discriminação. Os escores obtidos pelo Modelo 2 tiveram correlação alta com escores do Modelo 1 (0,9995) e com a NESSCA (0,9601). Com relação à informação total, houve perda de 1,08% de informação e o Alfa de Cronbach permaneceu muito próximo ao Alfa do Modelo 1.

A análise das Tabelas 1 e 2 sugere que o melhor modelo para representar a NESSCA através da TRI é o Modelo 5, uma vez que as correlações com NESSCA original e com escores do Modelo 1 foram altas, o Alfa de Cronbach se manteve alto e a perda de informação em relação ao Modelo 1 foi pequena. O Modelo 6 também apresenta correlações e Alfa altos, mas perde-se mais de 10% de informação.

Tabela 1: Modelos propostos a partir da retirada de itens com baixas discriminações

Modelo	Item retirado	Discriminação no modelo anterior	Correlação* com escores do Modelo 1	Correlação* com NESSCA	Informação do Modelo 1 sem este item	Perda de informação acumulada (%)	Alfa de Cronbach
2	NISTA	0,51	0,9995	0,9601	45,66	1,08	0,8107
3	RETR	0,64	0,9989	0,9583	45,53	2,45	0,8092
4	SENS	0,68	0,9925	0,9517	45,01	4,94	0,8108
5	FASC	0,67	0,9867	0,9452	45,42	6,54	0,8099
6	PIRAM	0,78	0,9908	0,9435	44,53	10,07	0,8286

* Todas as correlações foram significativas a 5% de nível de significância

Segunda etapa: validação

A calibração dos itens obtida para cada modelo foi aplicada nas observações dos 60 participantes do ensaio clínico para estimar seus comprometimentos no início do ensaio e realizar a validação através da comparação de correlações com outras escalas. As correlações dos escores TRI com os escores SARA, Barthel, BDI, SCAFI e CCFS estão na Tabela 2. As figuras 7 e 8 apresentam os gráficos de dispersão e de Bland-Altman dos escores do Modelo 5 versus NESSCA. Observamos que o gráfico de Bland-Altman do Modelo 5 ajustado no banco de validação foi semelhante ao gráfico do Modelo 1 ajustado no banco de calibração (Figura 5), com diferenças que chegaram a 4 pontos, sem padrões nas diferenças para maiores ou menores escores médios.

Tabela 2: Correlações* entre escores dos Modelos TRI e escalas NESSCA, SARA, Barthel, BDI, SCAFI e CCFS

Escore	NESSCA	SARA	Barthel	BDI	SCAFI	CCFS
NESSCA		0,620	-0,565	0,396	-0,537	0,513
Modelo Completo	0,924	0,744	-0,526	0,368	-0,581	0,552
Modelo 1	0,923	0,744	-0,520	0,364	-0,577	0,544
Modelo 2	0,923	0,742	-0,520	0,365	-0,577	0,543
Modelo 3	0,923	0,738	-0,522	0,361	-0,575	0,540
Modelo 4	0,916	0,706	-0,511	0,361	-0,575	0,541
Modelo 5	0,911	0,701	-0,507	0,356	-0,587	0,537
Modelo 6	0,903	0,734	-0,515	0,370	-0,589	0,532

* Todas as correlações foram significativas a 5% de nível de significância

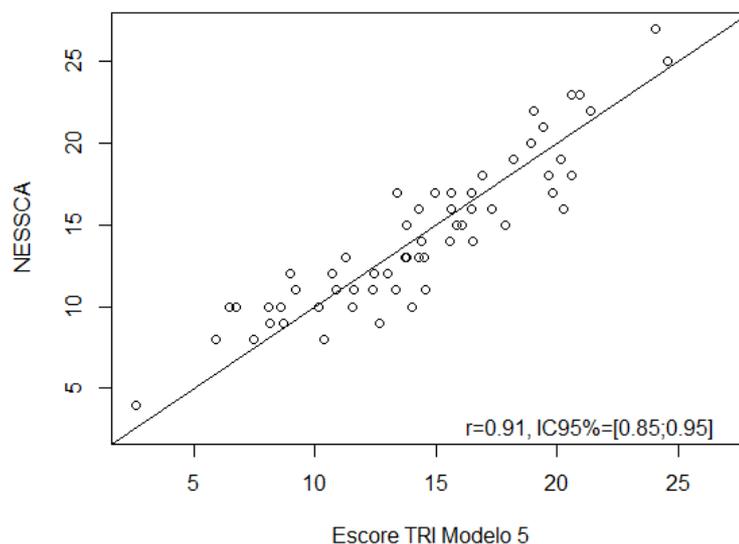


Figura 7: Diagrama de dispersão entre escores Modelo 5 e NESSCA

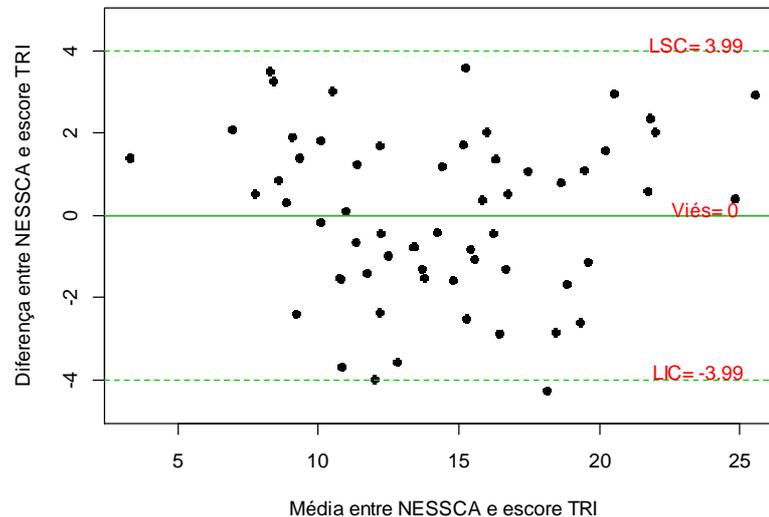


Figura 8: Gráfico de Bland-Altman para NESSCA e escore TRI Modelo 5

Discussão

A proposta de retirar os itens Câimbra e Vertigem, com base nos problemas de discriminação apresentados, corroborou a percepção do grupo de pesquisa que trabalhou na validação da NESSCA [13, 14] de que estes sintomas não discriminam os pacientes. Após esta alteração e agrupamentos de algumas categorias com base na curva característica dos itens, o Modelo 1 já apresentou algumas vantagens em relação ao Modelo Completo: aumento do Alfa de Cronbach, indicando maior consistência interna, com menos de 5% de perda de informação. Este resultado mostrou que a TRI pode ser utilizada para identificar itens e categorias que melhor explicam o comprometimento com DMJ.

A etapa de calibração mostrou que é possível retirarmos itens do Modelo 1 sem perder muita informação e as correlações com a NESSCA se mantêm altas. Os valores do Alfa de Cronbach mudam pouco com a retirada de itens, indicando que a consistência interna da escala não fica prejudicada com estas alterações. O modelo sugerido para estimar os escores TRI é o Modelo 5, que apresentou correlações altas com a NESSCA sem perder muita informação em relação ao Modelo 1. Isto significa que um modelo com 12 itens conseguiu representar tão bem o comprometimento com DMJ quanto o modelo completo de 18 itens, o que é uma vantagem. Além da redução do número de itens, devemos observar o ganho ao utilizar uma escala contínua para avaliar o traço latente. Com escores da TRI é possível diferenciar comprometimentos de pacientes que teriam o mesmo escore NESSCA, ou seja, não seria possível discriminá-los pela TCT.

A escala NESSCA se propõe a medir o comprometimento de ataxias (especialmente com DMJ) de uma forma mais abrangente. Por este motivo, na etapa de validação, ao compararmos correlações de Pearson entre escalas esperamos uma correlação alta, mas não perfeita, pois as escalas medem diferentes aspectos de um mesmo quadro clínico. Ao comparar com os resultados de outro estudo [13], a correlação com SARA e Barthel foram altas e significativas, conforme esperado. No entanto, em nosso estudo a correlação de NESSCA com BDI foi significativa igual a 0,396. As correlações entre escores TRI e as escalas SARA, Barthel, BDI, SCAFI e CCFS foram muito parecidas com as correlações destas escalas com a NESSCA. Este resultado valida os escores da

TRI, pois mostra que os escores TRI concordam com estas escalas da mesma forma que a NESSCA concorda.

Este estudo possui algumas limitações. Quanto ao número de populações envolvidas, com uma amostra de 106 sujeitos considerou-se impraticável detectar comportamento diferencial do item e admitiu-se que há apenas uma população. Com relação à quantidade de traços latentes, foram realizadas análises fatoriais exploratórias, e os resultados apontam indícios de que existem pelo menos 2 dimensões na escala NESSCA. No entanto, tendo em vista o tamanho de amostra e a grande quantidade de parâmetros para estimar em modelos multidimensionais, optou-se pela utilização de modelos unidimensionais [15].

Segundo alguns autores, para avaliar as propriedades de um questionário, não é preciso uma amostra de grande tamanho para ter uma imagem clara dos padrões de resposta [19,20]. No entanto, quando o objetivo é gerar escores da TRI precisos ou calibrar itens em um banco de dados, amostras maiores são necessárias [20]. Isto significa que neste estudo o tamanho de amostra pequeno foi um limitador não para a primeira etapa, em que sugerimos melhorias para a escala NESSCA, mas sim na segunda parte quando realizamos a validação ao aplicar a calibração dos itens em outro conjunto de dados. Por este motivo, é possível que as estimativas dos escores da TRI estejam pouco precisas. Além do tamanho de amostra, outro limitador é o fato de que todas as sugestões de melhorias para a NESSCA que se propõe aqui são válidas apenas quando avaliamos a DMJ, não sendo válidas para outras ataxias.

A aplicação da TRI ao escore NESSCA de avaliação da DMJ permitiu a identificação de itens problemáticos, que não estavam ajudando a medir o traço latente, e as categorias de alguns itens também puderam ser agrupadas. Apesar do tamanho de amostra ser um limitador, a calibração pareceu se comportar bem, pois quando aplicada em uma amostra diferente (etapa de validação) as correlações com as outras escalas fizeram sentido, inclusive os resultados foram semelhantes a estudos anteriores que utilizaram a abordagem clássica. A utilização da TRI em estudos futuros para avaliar o comprometimento com DMJ é recomendada, pois este método apresentou vantagens em relação ao método clássico, especialmente em relação ao menor número de itens necessários para medir o traço latente e o ganho em discriminação ao utilizar uma escala contínua.

Referências Bibliográficas

- [1] KOTZ, Samuel et al. (Org.). **Encyclopedia of statistical sciences**: International Statistical Review to Line Intersect Sampling. 2. ed. New Jersey: Wiley-Interscience, 2006.
- [2] ANJOS, Adilson Dos; ANDRADE, Dalton Francisco de. **Teoria da Resposta ao Item com o uso do R**. São Paulo, SP: ABE, 2012.
- [3] ANDRADE, D. F. de, TAVARES, H. R., VALLE, R. da C., **Teoria da Resposta ao Item: conceitos e aplicações**. São Paulo, SP: ABE, 2000.
- [4] PASQUALI, L.; PRIMI, R. (2003). **Fundamentos da Teoria da Resposta do Item – TRI**. Avaliação Psicológica, 2, 99-110.
- [5] EVERITT, Brian S.; HOWELL, David C. (Ed.). **Encyclopedia of Statistics in Behavioral Science**. Chichester: John Wiley & Sons Ltd, 2005.
- [6] ARAUJO, Eutalia Aparecida Candido de; ANDRADE, Dalton Francisco de; BORTOLOTTI, Silvana Ligia Vincenzi. Teoria da Resposta ao Item. **Rev Esc Enferm USP**, São Paulo, n. 43, p.1000-1008, 24 ago. 2009.
- [7] PASQUALI, L. **Psicometria**: teoria dos testes na psicologia e na educação. Petrópolis, RJ: Vozes, 2003.
- [8] REISE, Steven P. et al. Bifactor and Item Response Theory Analyses of Interviewer Report Scales of Cognitive Impairment in Schizophrenia. **Psychological Assessment**, vol. 23, n. 1, p.245-261, 2011.
- [9] VACCARINO, Anthony L et al. An Item Response Analysis of the Motor and Behavioral Subscales of the Unified Huntington’s Disease Rating Scale in Huntington Disease Gene Expansion Carriers. **Movement Disorders**, vol. 26, n. 5, p.877-884, 2011.
- [10] BALSIS, Steve et al. Gaining precision on the Alzheimer’s Disease Assessment Scale-cognitive: A comparison of item response theory-based scores and total scores. **Alzheimer’s & Dementia**, n. 8, p.288-294, 2012.
- [11] SAUTE, Jonas Alex Morales; JARDIM, Laura Bannach. Ataxias hereditárias. Rotinas em neurologia e neurocirurgia. Capítulo 26, p366-380.
- [12] HAUSER, Lisiane et al. Curva de Crescimento Usando Modelo Misto: uma aplicação na progressão da Doença De Machado-Joseph. **Revista HCPA**, vol. 29, n. 1, p.05-17, 2009.

[13] KIELING, C. A neurological examination score for the assessment of spinocerebellar ataxia 3 (SCA3). *European Journal of Neurology*, Apr. vol. 15(4) p.371-6, 2008.

[14] <http://www.clinicaltrials.gov/ct2/show/NCT01096082?term=sca&rank=12>, acessado em 17 de dezembro de 2012.

[15] CASTRO, STELA M. de J. Teoria da Resposta ao Item: aplicação na avaliação da intensidade de sintomas depressivos. 104 fls, Tese de Doutorado - Universidade Federal do Rio Grande do Sul. Porto Alegre, 28 de novembro de 2008.

[16] SAMEJIMA, F. Estimation of Latent Ability Using a Response Pattern of Graded Scores. **Psychometric Monograph**, n. 17. Richmond, VA: Psychometric Society, 1969.

[17] **R Development Core Team** (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

[18] Dimitris Rizopoulos (2006). ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses, *Journal of Statistical Software*, 17 (5), 1-25. URL <http://www.jstatsoft.org/v17/i05/>.

[19] COOK, K. F.,MICHAEL, E.I. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Qual Life Res* (2009) 18:447–460.

[20] ORLANDO, Maria. Critical Issues to Address when Applying Item Response Theory (IRT) Models. RAND Corp.

Anexos

Anexo 1: Escala NESSCA

Descrição dos itens da escala NESSCA completa [13].

Tabela 3: Descrição dos itens da NESSCA

Nome da Variável	Item	Gravidade	Categoria de resposta
CERAX	1 - Ataxia de marcha	Ausente	0
		Mínima: apenas ao andar na ponta dos pés, calcanhares, ou em conjunto	1
		Moderado: autonomia de marcha preservada	2
		Incapacidade de caminhar sem ajuda	3
		Cadeira de rodas ou acamados	4
CERAP	2 - Ataxia nos membros (bilateral)	Ausente	0
		Mínima: uma única prova alterado	1
		Moderado: duas provas alterados	2
		Importante: três provas alterados	3
		Provas: (a) dismetria, (b) rápido alternando os movimentos das mãos, e (c) de rebote do membro superior. Resultados positivos pode ser uni ou bilateral.	
NISTA	3 – Nistagmo	Ausente	0
		No olhar fixo; ou circular, depois de sacadas	1
		Permanente	2
OFTAP	4 -Oftalmoplegia externa progressiva	Ausente	0
		Supranuclear: síndrome médio de fascículo longitudinal; ou limitação no olhar para cima ou para a convergência	1
		Oftalmoplegia nuclear, com estrabismo	2
PIRAM	5 - Achados piramidais	Ausente	0
		Poucos reflexos rápidos	1
		Geral hiperreflexia, ou clônus, ou sinal de Babinski	2
		Três conclusões: (a) hiperreflexia geral, (b) a espasticidade, (c) clônus, (d) sinal de Babinski, (e) paresia	3
		Quatro ou cinco dos sinais acima mencionados	4
DISAR	6 – Disartria	Ausente	0

		Leve: Dificuldade de fala, mas fácil de entender	1
		Moderado: discurso compreensível, mas com dificuldade	2
		Grave: discurso de difícil compreensão	3
		Anartria	4
		Ausente	0
DISF	7 – Disfagia	Leve	1
		Importante: ocorrendo todos os dias	2
		Ausente	0
FASC	8 – Fasciculações	Fasciculação contração no rosto	1
		Difusa, ou em outras partes do corpo	2
		Ausente	0
		Uma prova alterada: Redução de (a) ou (b) ou (c): 2-4 erros, em média, em ambos os pés	1
SENS	9 - Perda sensorial	Duas provas alteradas	2
		Perda total do sentido de vibração nos dedos do pé, ou 5 ou mais erros em uma das provas de discriminação, ou três provas alteradas	3
		Ausente	0
		Leve, acionado por movimentos voluntários	1
DISTO	10 – Distonia	Moderados, prejudicando, em algum grau, movimentos voluntários	2
		Quase constantes, prejudicando severamente os movimentos voluntários	3
		Ausente	0
		Moderada: não impede mobilização, passiva total	1
RIGID	11 – Rigidez	Importante: para impedir a mobilização, passiva total	2
		Ausente	0
BRADI	12 – Bradicinesia	Movimentos lentos, com redução da amplitude	1
		Movimentos dificilmente podem ser feitos	2
		Ausente	0
RETR	13 - Retração palpebral	Presente	1
		Ausente	0
BLEFARO	14 – Blefarospasmo	Presente	1
		Ausente	0
AMIO	15 - Amiotrofia distal	Presente	1
		Normal	0
ESFI	16 - Função do Esfíncter	Urgência	1
		Incontinência	2
		Ausente	0
CAIM	17 – Câimbra	Presente	1
		Ausente	0
VERT	18 – Vertigem	Presente	1

Anexo 2: Gráficos do Modelo TRI Completo

Figuras com as curvas características dos itens e informação total do Modelo TRI Completo (18 itens, com categorias originais de resposta da escala NESSCA)

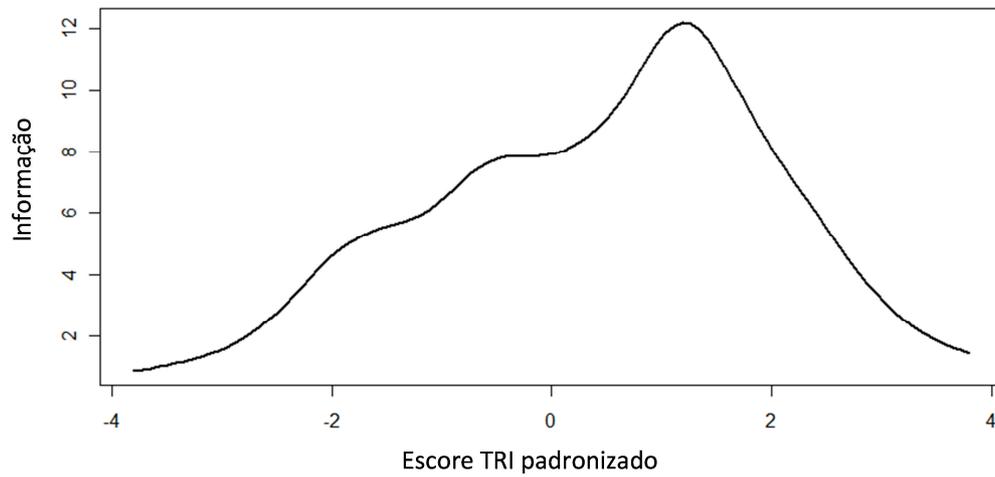


Figura 9: Função de Informação do Item Modelo Completo

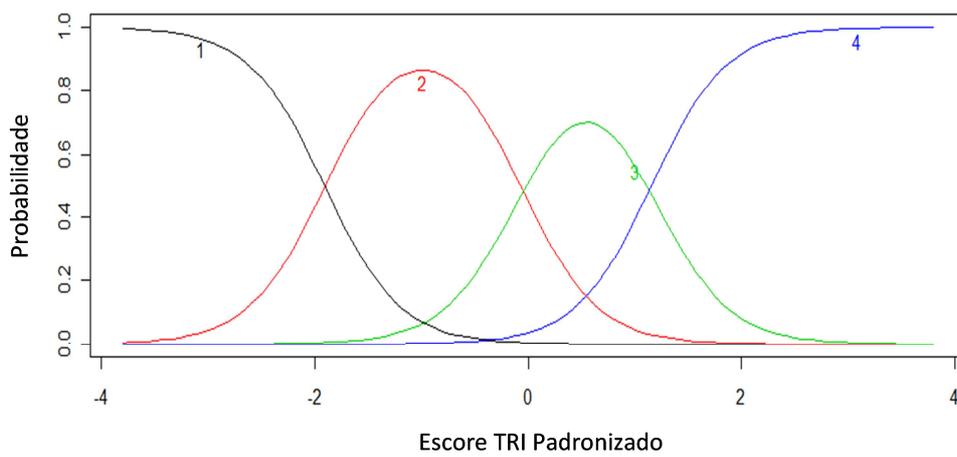


Figura 10: Curva Característica do Item CERAX, Modelo Completo

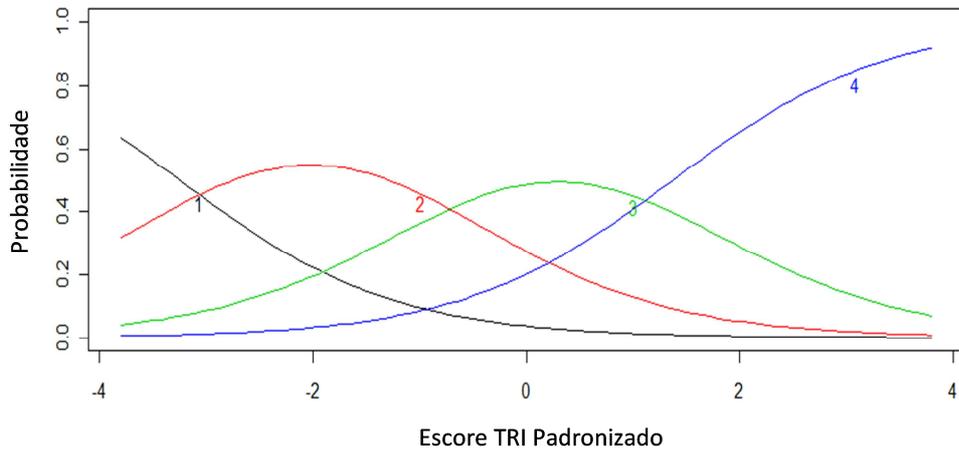


Figura 11: Curva Característica do Item CERAP, Modelo Completo

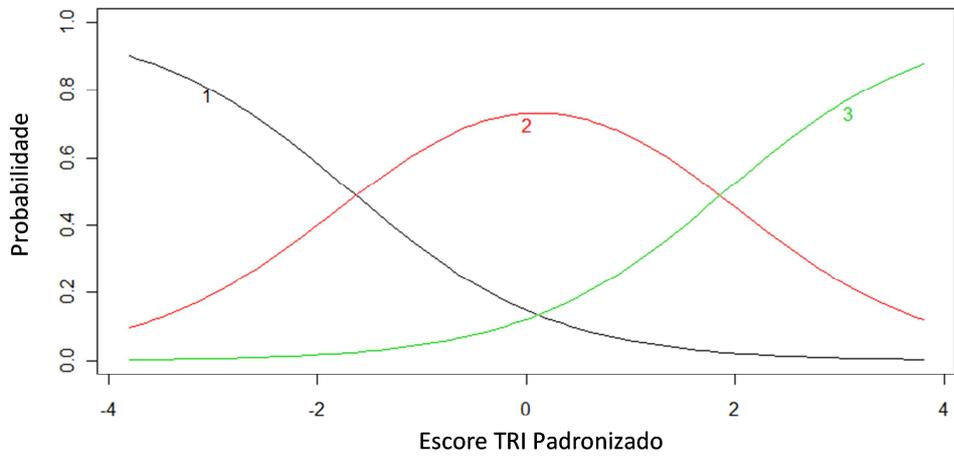


Figura 122: Curva Característica do Item OFTAP, Modelo Completo

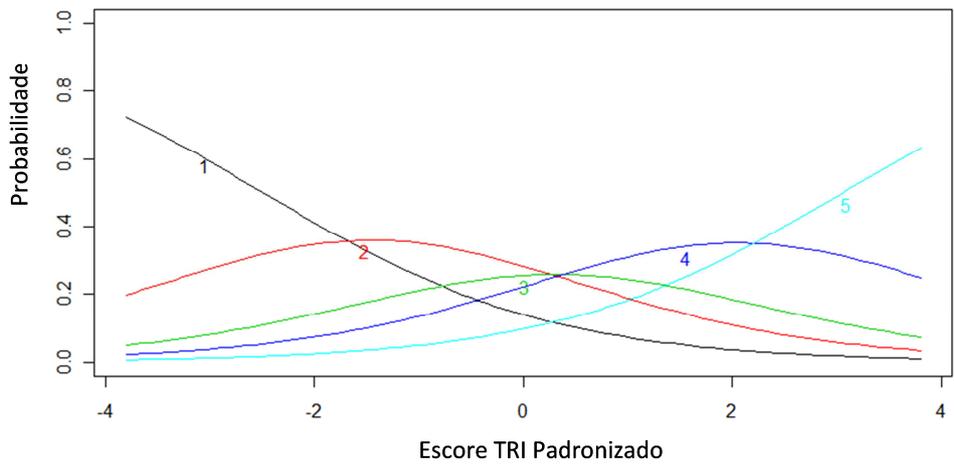


Figura 13: Curva Característica do Item PIRAM, Modelo Completo

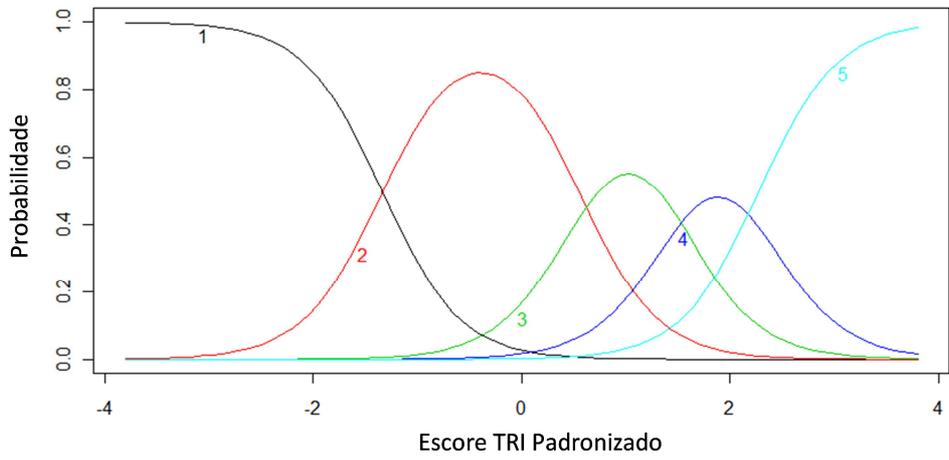


Figura 14: Curva Característica do Item DISAR, Modelo Completo

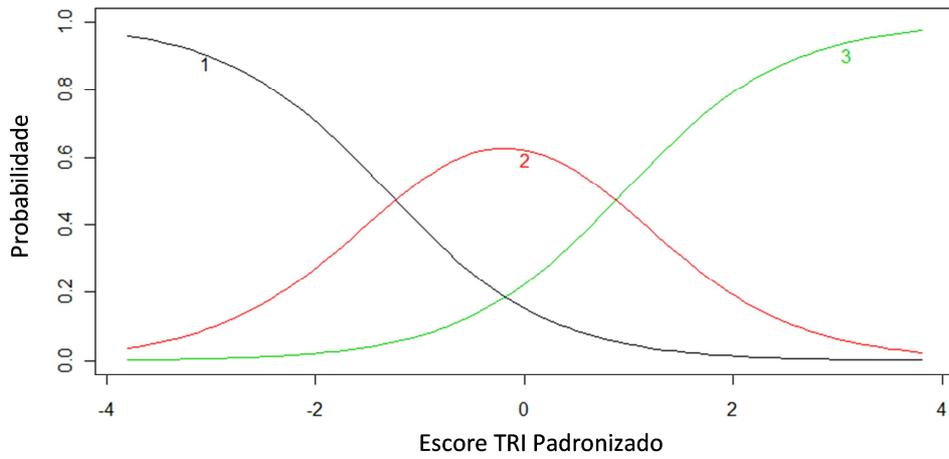


Figura 15: Curva Característica do Item DISF, Modelo Completo

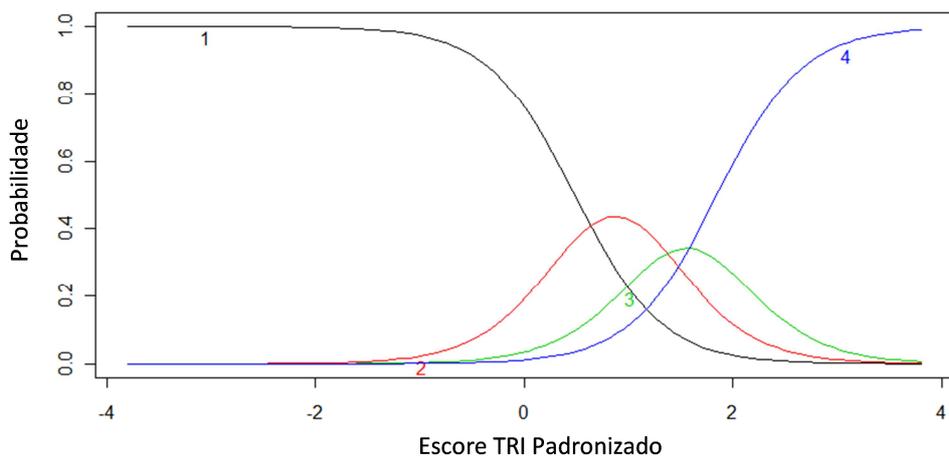


Figura 16: Curva Característica do Item DISTO, Modelo Completo

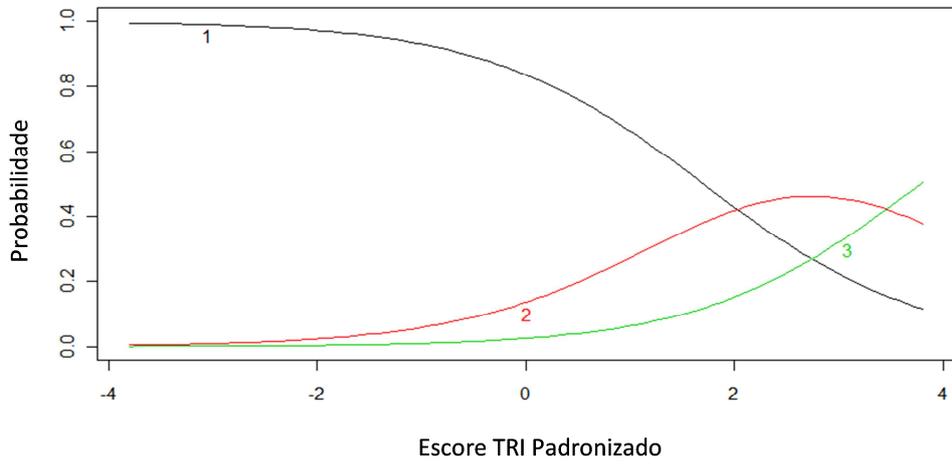


Figura 17: Curva Característica do Item RIGID, Modelo Completo

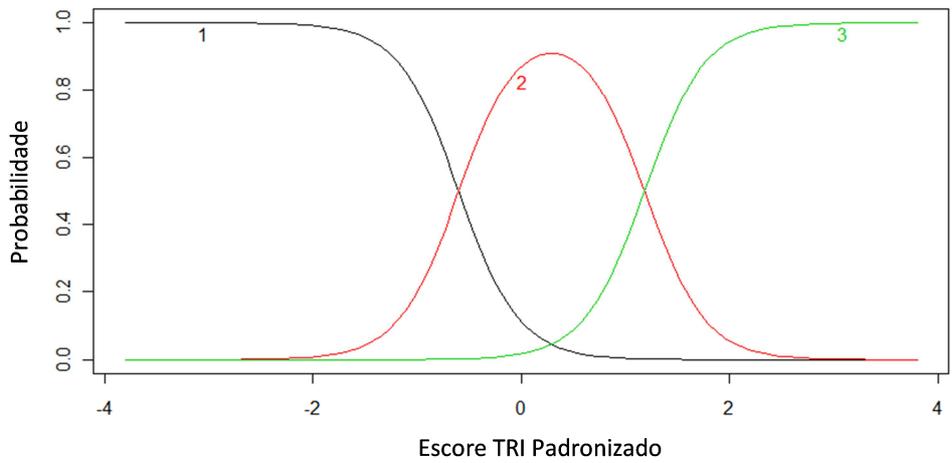


Figura 18: Curva Característica do Item BRADI, Modelo Completo

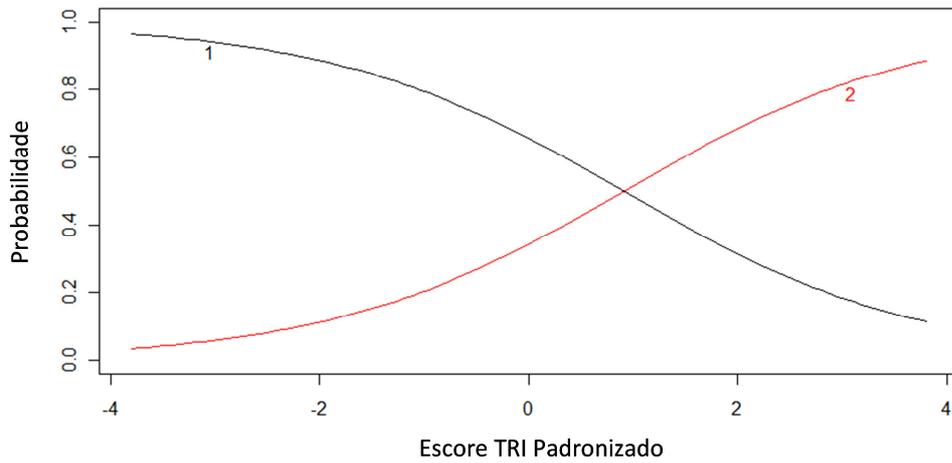


Figura 19: Curva Característica do Item RETR, Modelo Completo

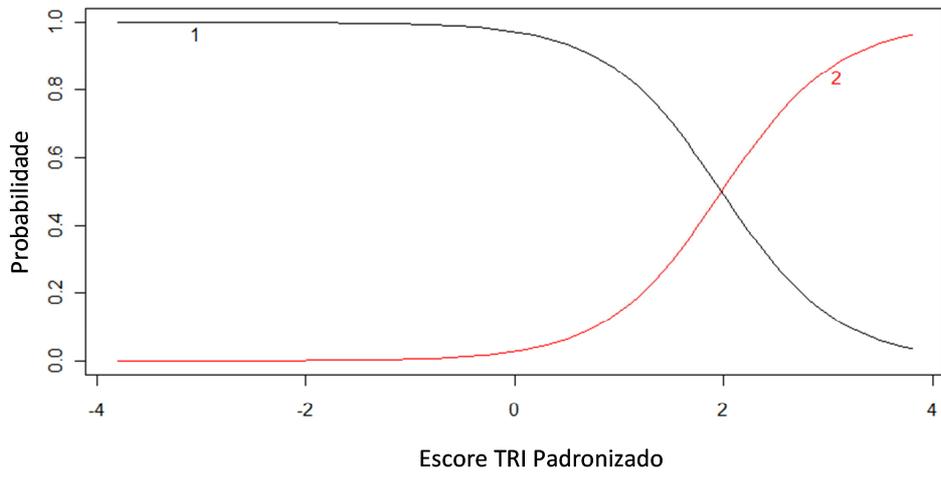


Figura 20: Curva Característica do Item BLEFARO, Modelo Completo

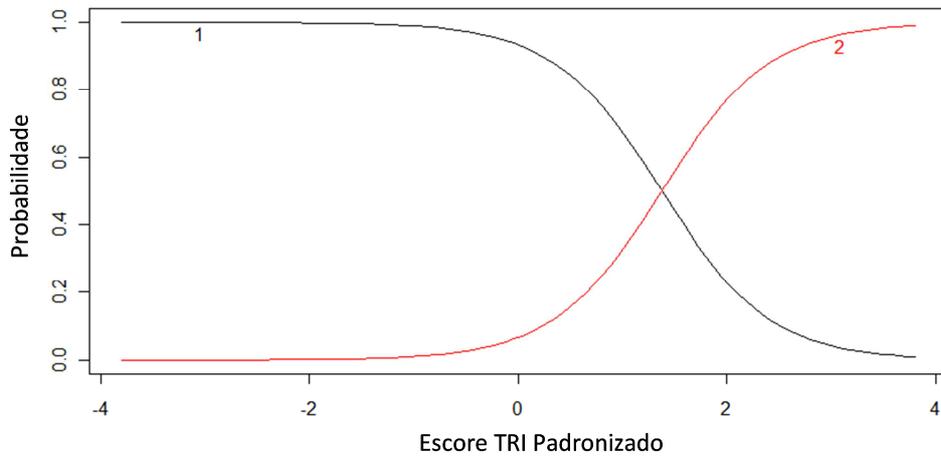


Figura 21: Curva Característica do Item AMIO, Modelo Completo

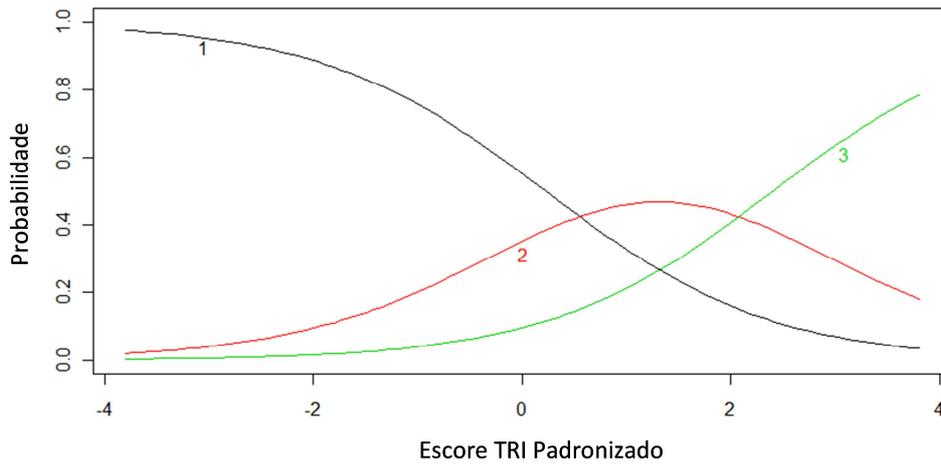


Figura 22: Curva Característica do Item ESFI, Modelo Completo

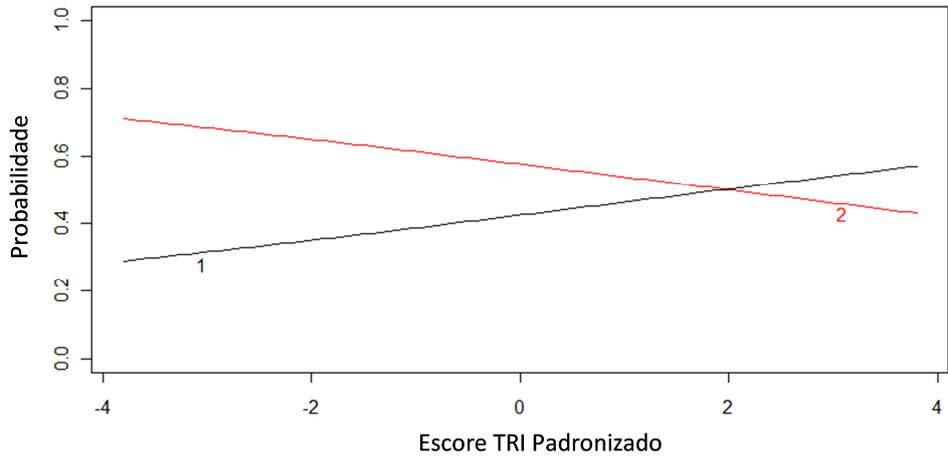


Figura 23: Curva Característica do Item CAIM, Modelo Completo

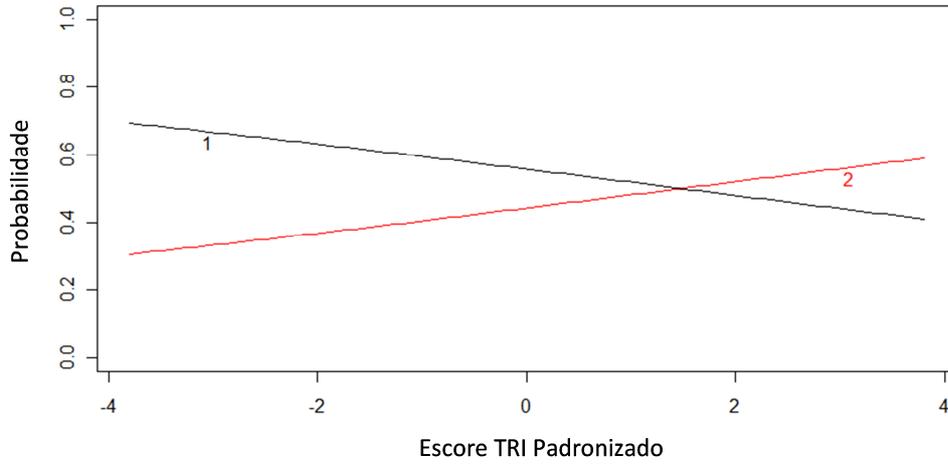


Figura 24: Curva Característica do Item VERT, Modelo Completo

Anexo 3: Estimativas pontuais dos parâmetros dos modelos TRI ajustados

Tabela 4: Estimativas pontuais dos parâmetros do Modelo Completo

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,484	0,129	1,075		4,086
CERAP	2,655	0,552	1,097		1,210
NISTA	1,430	4,303			0,692
OFTAP	1,382	2,068			1,083
PIRAM	2,661	0,444	1,133	2,998	0,711
DISAR	1,322	0,661	1,672	2,387	2,329
DISF	1,052	0,830			1,618
FASC	0,387	4,266			0,666
SENS	2,110	0,070	1,287		0,805
DISTO	0,554	1,419	2,317		2,099
RIGID	1,698	3,634			1,013
BRADI	0,621	1,495			2,052
RETR	0,778				0,644
BLEFARO	2,071				1,821
AMIO	1,565				2,024
ESFI	0,426	2,539			1,034
CAIM	0,544				-0,218
VERT	14,475				0,024

Tabela 5: Estimativas pontuais dos parâmetros do Modelo 1

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,485	0,129	1,075		3,999
CERAP	2,662	0,553	1,097		1,206
NISTA	1,844				0,507
OFTAP	1,376	2,057			1,09
PIRAM	2,648	0,442	1,127	2,98	0,715
DISAR	1,311	0,655	1,659	2,364	2,276
DISF	1,041	0,821			1,647
FASC	0,332				0,763
SENS	1,87	0,009			0,966
DISTO	0,554	1,423	2,321		2,088
RIGID	1,676	3,585			1,028
BRADI	0,621	1,491			2,055
RETR	0,779				0,641
BLEFARO	2,068				1,82
AMIO	1,594				1,937
ESFI	0,442	2,517			1,046

Tabela 6: Estimativas pontuais dos parâmetros do Modelo 2

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,492	0,120	1,071		3,945
CERAP	1,688	0,566	1,098		1,191
OFTAP	1,380	2,047			1,093
PIRAM	2,636	0,444	1,114	2,953	0,721
DISAR	1,311	0,644	1,644	2,350	2,406
DISF	1,046	0,812			1,653
FASC	0,327				0,758
SENS	1,868	0,001			0,956
DISTO	0,544	1,413	2,310		2,102
RIGID	1,670	3,538			1,027
BRADI	0,630	1,487			2,048
RETR	0,768				0,645
BLEFARO	2,064				1,819
AMIO	1,596				1,918
ESFI	0,415	2,513			1,043

Tabela 7: Estimativas pontuais dos parâmetros do Modelo 3

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,508	0,117	1,070		3,846
CERAP	2,679	0,568	1,084		1,204
OFTAP	1,380	2,029			1,101
PIRAM	2,637	0,457	1,120	2,981	0,712
DISAR	1,321	0,638	1,634	2,341	2,392
DISF	1,059	0,810			1,633
FASC	0,325				0,747
SENS	1,855	0,005			0,681
DISTO	0,535	1,394	2,278		2,146
RIGID	1,666	3,576			1,025
BRADI	0,630	1,466			2,091
BLEFARO	2,027				1,867
AMIO	1,585				1,925
ESFI	0,406	2,479			1,058

Tabela 8: Estimativas pontuais dos parâmetros do Modelo 4

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,585	0,143	1,183		3,082
CERAP	2,883	0,594	1,210		1,073
OFTAP	1,408	2,135			1,053
PIRAM	2,506	0,407	1,092	2,849	0,764
DISAR	1,289	0,656	1,667	2,357	2,564
DISF	1,607	0,863			1,571
FASC	0,388				0,670
DISTO	0,538	1,375	2,229		2,456
RIGID	1,698	3,612			1,029
BRADI	0,619	1,523			2,085
BLEFARO	2,166				1,733
AMIO	1,741				1,695
ESFI	0,437	2,550			1,044

Tabela 9: Estimativas pontuais dos parâmetros do Modelo 5

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,621	0,148	1,219		1,842
CERAP	2,985	0,612	1,254		1,025
OFTAP	1,459	2,212			1,004
PIRAM	2,472	0,402	1,081	2,816	0,777
DISAR	1,287	0,658	1,671	2,352	2,585
DISF	1,092	0,888			1,504
DISTO	0,529	1,347	2,182		2,647
RIGID	1,681	3,536			1,049
BRADI	0,605	1,497			2,206
BLEFARO	2,136				1,796
AMIO	1,786				1,625
ESFI	0,432	2,509			1,071

Tabela 10: Estimativas pontuais dos parâmetros do Modelo 6

Item	β_{i0}	β_{i1}	β_{i2}	β_{i3}	β_i
CERAX	1,575	0,125	1,174		3,123
CERAP	2,812	0,586	1,167		1,118
OFTAP	1,440	2,518			1,031
DISAR	1,320	0,657	1,694	2,401	2,393
DISF	1,079	0,857			1,557
DISTO	0,535	1,387	2,252		2,353
RIGID	1,670	3,562			1,042
BRADI	0,617	1,490			2,177
BLEFARO	2,115				1,808
AMIO	1,706				1,750
ESFI	0,409	2,448			1,096