

Heitor Braga de Melo

**Análise de Componentes Principais e Análise Fatorial de
dados de Fadiga**

Uberlândia

Dezembro, 2016

Heitor Braga de Melo

Análise de Componentes Principais e Análise Fatorial de dados de Fadiga

Trabalho de Conclusão de Curso apresentado à Faculdade de Matemática da Universidade Federal de Uberlândia como requisito parcial para a obtenção do grau de Bacharel em Estatística.

Universidade Federal de Uberlândia - UFU

Faculdade de Matemática

Bacharelado em Estatística

Orientador: Tiago Moreira Vargas

Uberlândia

Dezembro, 2016

Heitor Braga de Melo

Análise de Componentes Principais e Análise Fatorial de dados de Fadiga

Trabalho de Conclusão de Curso apresentado à Faculdade de Matemática da Universidade Federal de Uberlândia como requisito parcial para a obtenção do grau de Bacharel em Estatística.

Trabalho aprovado. Uberlândia, 02 de Dezembro de 2016:

Dr. Tiago Moreira Vargas
Orientador

Msc. Patrícia Viana da Silva
Convidada

Dra. Raquel Romes Linhares
Convidada

Uberlândia
Dezembro, 2016

Agradecimentos

Os agradecimentos principais vão aos meus magníficos professores doutores Priscila Neves faria, Rogério de Melo Costa Pinto, Raquel Romes Linhares e Tiago Moreira Vargas. Obrigado pelo incentivo e todo apoio que me deram durante todos esses anos. Um agradecimento a mestre Patrícia Viana pelas brigas e puxões de orelha, que me fizeram crescer na minha vida acadêmica.

Agradeço aos meus pais Adolfo e Silvânia e a minha irmã Ana Paula por todo carinho, pois vocês são os alicerces de minha vida.

Um agradecimento especial a 5º Tuma de Estatística que juntos conseguimos conciliar os trabalhos, as provas e a nossa vida pessoal. Vocês ficarão eternamente em meu coração.

Agradeço aos membros do grupo de pesquisa Promis – Bruno Simão Teixeira, Dr. Carlos Henrique Martins da Silva, Tânia Mendonça Marques.

Resumo

O objetivo do trabalho é o estudo das técnicas multivariadas de análise de componentes principais e análise fatorial através do método de componentes principais utilizando dados obtidos do questionário Fadiga-Promis. Os dados foram coletados de pacientes acompanhados no ambulatório do Hospital das Clínicas da Universidade Federal de Uberlândia (HCU) mediante assinatura de termo de consentimento dos mesmos, e aprovação do Comitê de Ética Humana da UFU. A análise foi conduzida a partir de um questionário de 82 itens aplicados a 1000 pacientes com idade igual ou superior a 21 anos dividido em dimensões distintas.

Palavras-chaves: Fadiga. Técnicas multivariadas.

Abstract

The objective this work is the study of the multivariate techniques of principal component analysis and factor analysis through principal component method in data obtained Fadiga-Promis questionnaire. The data were collected from patients of the Federal University of Uberlândia Hospital by signing their consent, and approval of the UFU Human Ethics Committee. The analysis was based on a questionnaire of 82 items applied to 1000 patients aged 21 years.

Palavras-chaves: Fadiga. Multivariate techniques.

Sumário

1.Introdução.....	4
2.Referencial Teórico	7
2.1.Componentes Principais	7
2.1.2.Critérios de escolha do número de componentes	8
2.1.3.Método Biplot	10
2.2.Análise Fatorial	10
3.Aplicações	13
4.Conclusão	28
5.Referências	29
6.Apêndice	32

1. Introdução

A fadiga é um dos sintomas mais relatados pelos pacientes com doenças crônicas (YORKSTON et al., 2010) levando-os à limitações sociais como: à capacidade de realizar atividades de vida diária (AVD's) e a trabalhar de forma eficaz, à diminuição da auto-estima e da qualidade de vida (QV) (PROMIS COOPERATIVE GROUP, 2008; MENDES et al., 2008; MOTA et al., 2010; PROMIS – Item definition - PROMIS Cooperative Group, 2010; ASTUDILLA et al., 2011). É um sintoma subjetivo que impacta os aspectos físico, mental e social devido à perda de energia ou mesmo sensação de exaustão e possui características diferentes daquelas observadas em sintomas de depressão ou fraqueza muscular (KRUPP et al., 1989; PROMIS COOPERATIVE GROUP, 2008), sendo mais frequente entre as mulheres e idosos (MOTA et al., 2010).

Profissionais da saúde utilizam questionários para avaliar o sintoma da fadiga no seu estado inicial e os resultados do tratamento. Estes testes, geralmente fornecem uma maneira conveniente de avaliar como o tratamento tem resultado na dor, na incapacidade e na qualidade de vida relacionada a saúde (QVRS) (MAHER et al., 2007). Com o avanço tecnológico, o Instituto Nacional de Saúde Americano (NIH) lançou o *Patient Reported Outcome Measurement Information System* (PROMIS) que surge como uma nova promessa de avaliação da qualidade de vida relacionada à saúde, com um de seus bancos de itens que avalia exclusivamente a fadiga e que irá complementar e revolucionar as formas de avaliações existentes (FRIES et al., 2005). É um banco de itens que garante precisão com um número menor de pacientes para análise estatística (FRIES et al., 2005) por se utilizar de modelos da Teoria de Resposta ao Item (TRI) e o Teste Adaptativo Computadorizado (CAT) (WARE et al., 2004). Com a TRI e o CAT, o entrevistado só precisa responder a um pequeno número de elementos informativos para que se possa estimar com precisão o que teria sido obtido se tivesse respondido a todo o conjunto de itens (CELLA et al., 2007). Como a aplicação e interpretação precisa dos resultados, o profissional será capaz de escolher o melhor plano de intervenção no tratamento de cada paciente e modificá-lo, dependendo das respostas dadas por estes pacientes (CELLA et al., 2007).

Como já se sabe, o PROMIS foi criado por pesquisadores norte-americanos e, portanto, criado de acordo com a cultura deste país (PROMIS Cooperative Group, 2008). Esse novo sistema de avaliação teve como desafio o desenvolvimento de um grande banco

de itens retirados de instrumentos clássicos e concentrados em domínios de avaliação física e psicossocial para serem aplicados por meio do PROMIS® (PROMIS® Cooperative Group, 2008; CELLA et al., 2004; CELLA et al., 2007a; CELLA et al., 2007b).

A Saúde Física é um dos domínios contemplados neste amplo banco de itens e é representada por dois subdomínios, a função física e Sintomas. No subdomínio Sintomas está inserido o banco de itens Fadiga (CELLA et al., 2007b; PROMIS®CooperativeGroup, 2008). A fadiga é um sintoma de difícil definição e, portanto, ainda sem consenso entre os pesquisadores. Existe uma concordância de que a fadiga tem etiologia multifatorial e é um sintoma subjetivo de sensação desagradável de cansaço que dificilmente alivia com medidas usuais para restauração de energia, e que promove impacto negativo nos aspectos físicos, psíquicos e emocionais das pessoas comprometidas, ou seja, em sua qualidade de vida relacionada à saúde (QVRS). A duração e intensidade deste sintoma são variáveis entre as pessoas e reduz em diferentes graus, sua habilidade para executar as atividades de vida diária, além de reduzir a capacidade profissional e de socialização (MOTA et al., 2005; MOTA e PIMENTA, 2006; PROMIS® – Item definition – PROMIS®CooperativeGroup, 2008).

Pessoas com a sintomatologia de fadiga em decorrência de uma ou mais doenças crônicas, apresentam prejuízo físico e psicossocial com consequente impacto negativo na QVRS. A avaliação da influência da fadiga na QVRS do paciente tem sido realizada por meio de instrumentos genéricos ou de escalas específicas para avaliação da fadiga. Dentre as escalas específicas mais utilizadas e traduzidas, adaptadas e validadas para a população brasileira estão a Escala de Severidade da Fadiga (FSS) (KRUPP et al., 1989; MENDES et al., 2008), a Escala de Impacto de Fadiga (FIS) (FISK et al., 1994) modificada para Escala de Impacto de Fadiga Modificada (MIFS) (KOS et al., 2006; PAVAN et al., 2007). Para utilizar estes instrumentos, sejam eles genéricos ou específicos, os pacientes além de responder a todas as suas questões, o que demanda tempo, ainda existem as limitações metodológicas desses instrumentos de medida. O PROMIS® traz a vantagem de o paciente responder a um número mínimo de itens, com o uso da TRI e do CAT além de resultar em respostas e informações mais precisas (PROMIS®CooperativeGroup, 2008).

Por isso, é necessário que os investigadores adaptem o questionário original para que seja compreensível e relevante para a nova população, pois é um método que trará benefícios para avaliar o sintoma fadiga. Este processo é chamado adaptação transcultural (MAHER et al., 2007), e permite aos pesquisadores testarem, modificarem e estenderem

teorias atuais de QV às outras populações (PEÑA, 2007). Esse novo sistema de avaliação teve como desafio o desenvolvimento de um grande banco de itens retirados de instrumentos clássicos e concentrados em domínios de avaliação física e psicossocial para serem aplicados por meio do PROMIS® (PROMIS® Cooperative Group, 2008; CELLA et al., 2004; CELLA et al., 2007a; CELLA et al., 2007b).

Diante do exposto, percebe-se a importância da pesquisa em torno dessa temática, no intuito de descrever para os especialistas a real QV dos pacientes, contribuindo, assim, para o redirecionamento das estratégias e ações que fomentam o controle à doença, visto que ainda não há uma forma de avaliação eficaz no país.

Além disso, é necessária também, a validação das propriedades psicométricas deste banco de itens para a população brasileira e demais populações que desejarem utilizar este banco de itens em seu país. Por isso, para se obter dados com qualidade científica, é imprescindível a realização de testes das propriedades de medida desses instrumentos por meio de múltiplas análises estatísticas (CICONELLI et al., 1999).

As propriedades de medida (ou psicométricas) podem variar de maneira considerável para cada população (McHORNEY et al., 1994; MCCARTHY et al., 2002). Daí a necessidade de se testar estatisticamente tais parâmetros psicométricos dentro de um contexto cultural específico e na condição de interesse. Somente dessa forma é possível obter um bom instrumento: válido, confiável e sensível.

A validade verifica se o instrumento realmente mede o conceito alvo, e não outro (GUYATT; FEENY; PATRICK, 1993). O estabelecimento da validade de um instrumento recém formulado, geralmente é feito por comparação com outro instrumento (padrão ouro) já aceito e que mede o mesmo conceito (GUYATT; FEENY; PATRICK, 1993; EISER; MORSE, 2001).

A confiabilidade diz respeito à precisão do instrumento, ou seja, é a estimativa da quantidade de “erro” inerente ao processo de aferição (GRIEP et al., 2003). Em geral, indica se pacientes estáveis possuem resultados semelhantes após várias administrações (EISER; MORSE, 2001).

A análise de componentes principais reduz a dimensão e descreve a variabilidade original do meu instrumento em termos de um número menor de variáveis, chamadas de fatores comuns (MINGOTI, et al., 2005).

A análise fatorial através do método de componentes principais calcula índices que representam o instrumento através das variáveis originais, chamados de escores, que

serão utilizados para a interpretação das cargas fatoriais referentes a cada índice (MINGOTI, et al., 2005).

O objetivo do trabalho é avaliar as propriedades psicométricas do banco de itens Fadiga-Promis para avaliar a qualidade dos itens através de técnicas multivariadas.

2. Referencial Teórico

2.1 Componentes Principais

A análise de componentes principais é uma técnica multivariada que tem como objetivo a redução de massa dos dados, com a menor perda possível da informação. Consiste em transformar um grupo de variáveis originais em outro grupo de variáveis da mesma dimensão denominados de componentes principais. Os componentes principais possuem propriedades importantes, pois são uma combinação linear de todas as variáveis originais, são independentes entre si e são estimados com o propósito de reter a maior informação possível referente a variabilidade total dos dados. A variabilidade observada é redistribuída com o intuito de se obter eixos ortogonais não correlacionados.

Segundo STRAPASSON (2000) os componentes principais consistem em transformar um conjunto de m variáveis originais $X_1, X_2, \dots, X_m$, pertencentes a n indivíduos, em um novo conjunto de variáveis $Y_1, Y_2, \dots, Y_m$, de dimensão igual, denominadas componentes principais. Cada componente principal é a combinação linear das variáveis originais, construído para explicar a variabilidade total das variáveis originais e não correlacionadas entre si.

O primeiro componente principal é o responsável por explicar a maior variabilidade das variáveis originais. Segundo Morrison (1976) o primeiro componente principal (Y_1) de um grupo de m variáveis, $X_1, X_2, \dots, X_m$, contidas em uma matriz $X' = (X_1, X_2, \dots, X_m)$ é definido como uma combinação linear:

$$Y_1 = a_{11}X_1 + a_{21}X_2 + \dots + a_{m1}X_m = a'_1X$$

Onde os coeficientes a_{i1} são elementos do autovetor a_1 , associado ao maior autovalor (λ_1) da matriz da covariância amostral, H , das variáveis X_{is} :

$$H = \begin{bmatrix} \widehat{Var}(X_1) & \widehat{Cov}(X_1X_2) \dots & \widehat{Cov}(X_1X_m) \\ \widehat{Cov}(X_2X_1) & \ddots & \widehat{Cov}(X_2X_m) \\ \widehat{Cov}(X_mX_1) & \widehat{Cov}(X_pX_2) \dots & \widehat{Var}(X_m) \end{bmatrix}$$

O segundo componente principal (Y_2) está associado ao segundo maior autovalor (λ_2). Com isso os demais componentes serão associados de forma análoga até que toda a variabilidade tenha sido explicada. As raízes características ou os autovalores ordenados, ou seja, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$, são as variâncias amostrais dos componentes principais.

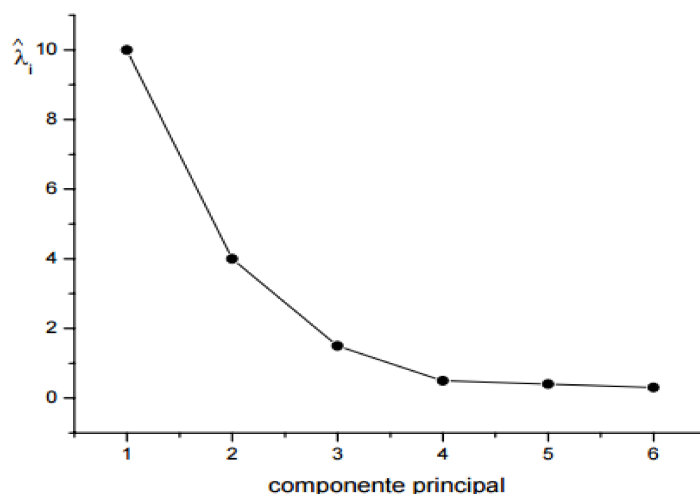
2.1.2 Critérios de escolha do número de componentes

Existe uma subjetividade na escolha dos componentes a serem retidos. Não há uma resposta definitiva para essa questão. Devido a essas dificuldades para definir o número de componentes principais que deve ser utilizado na análise dos dados experimentais, surgiram critérios que auxiliaram na escolha e facilitaram a interpretação dos dados. Os critérios para a definição envolvem a quantidade da variação amostral explicada, o tamanho relativo dos autovalores e a interpretação subjetiva dos componentes.

Critério do “Scree Plot”

Uma ferramenta visual importante para a definição do número de componentes é o gráfico “scree plot” (Cattell, 1996), que mostra os valores numéricos dos autovalores $\hat{\lambda}_i$ de acordo com a respectiva ordem i . O termo “scree” está relacionado ao acúmulo de rochas nas bases de um penhasco, portanto os “scree plots” são “gráficos de cotovelos”. Baseia-se na magnitude dos autovalores, pois quando eles começam a se tornar “constantes” utilizamos o critério para a seleção, ou seja, tomamos os autovalores acima das magnitudes constantes dos outros autovalores. Na Figura 1 observa-se que é formado um cotovelo na posição $i=4$, com isso significa que os componentes acima de $\hat{\lambda}_3$ possuem aproximadamente a mesma magnitude e são relativamente pequenos, indicando assim que os três primeiros componentes são suficientes para resumir a variação amostral total.

Figura 1: “Scree Plot” de um exemplo com $p=6$ componentes principais para ilustrar o processo de determinação do número de componentes apropriados a serem retidos.



Critério de Kaiser

O critério de Kaiser é um método exploratório e o mais utilizado entre todos. Descreve o primeiro componente principal, λ_1 , como um eixo que minimiza a soma das distâncias euclidianas ao quadrado de todos os pontos até aquele eixo. O segundo componente, λ_2 , corresponde a projeção ortogonal ao primeiro componente, que retém a segunda maior variância dos dados amostrais. O terceiro componente principal, λ_3 , é um vetor ortogonal aos dois primeiros e que tem o terceiro posto no ranking, sendo assim a terceira maior variância retida, e assim sucessivamente para os outros componentes.

Segundo Artes (1998), considera-se que o número de fatores deve ser igual ao número de autovalores maiores ou iguais à médias das variâncias das variáveis analisadas. Quando se utiliza a matriz de correlação (variáveis padronizadas) esse critério corresponde à exclusão dos componentes com autovalores inferiores a um.

Para Hair (2005) o método de Kaiser baseia-se no critério de que qualquer autovalor deve explicar a variância de pelo menos uma variável para que o mesmo seja mantido para a interpretação. Portanto, cada variável contribui com valor (1,0) do autovalor total, logo, apenas autovalores maiores que um são considerados e aqueles autovalores menores que um são descartados.

2.1.3 Método Biplot

Um método gráfico bastante utilizado é o Biplot desenvolvido por Gabriel (1971) que permite representar graficamente resultados de ACP (análise de componentes principais). Tem como objetivo utilizar marcadores de dados de linha e coluna que formam uma tabela de dupla entrada, sendo eles calculados a partir da decomposição por valores singulares da matriz de dados. No Biplot é possível visualizar as relações existentes entre as variáveis, entre observações e entre variáveis e observações.

2.2 Análise fatorial

A análise fatorial foi criada por Spearman em 1904, de mesmo modo que os componentes principais tem como objetivo principal descrever a variabilidade original do matriz aleatório X , em termos de um número menor m de variáveis aleatórias, mais conhecidas como fatores comuns que estão relacionadas com o matriz original X através de um modelo linear. Com isso parte da variabilidade de X está atribuída aos fatores comuns e o restante da variabilidade está atribuído aos fatores específicos ou erros.

Geralmente utilizada quando se tem um número grande de variáveis correlacionadas, a análise fatorial identifica um número menor de variáveis alternativas, não correlacionadas e que, de algum modo, sumarizam as informações principais das variáveis originais, chamadas de fatores ou variáveis latentes.

Segundo Rencher (2001), a análise fatorial é expressa por variáveis originais que são mais conhecidas como combinações lineares de fatores observáveis que busca explicar a covariância ou correlação entre as variáveis. Já para Härdle e Simar (2003) a análise fatorial é o pós-rotação que fornece a máxima interpretação e envolve grandes variedades de procedimentos numéricos.

A análise fatorial é dividida em dois tipos: exploratória e confirmatória. A exploratória como o próprio nome diz, busca encontrar os fatores subjacentes às variáveis originais amostradas, com isso o pesquisador utiliza essa técnica devido ele não ter a noção clara de quantos fatores fazem parte do modelo e nem o que esses representam. Já na análise confirmatória o pesquisador tem em mãos um modelo fatorial pré-especificado que deseja-se verificar se é consistente com os dados amostrais.

Modelo análise fatorial através da matriz de correlação

Seja $X_{p \times 1}$ um vetor aleatório com vetor de médias μ , onde $\mu = (\mu_1, \mu_2 \dots \mu_p)'$, matriz de covariâncias $\Sigma_{p \times p}$ e matriz de correlação $P_{p \times p}$. Sejam as variáveis originais

padronizadas onde $\mu_i e \sigma_i$ representam respectivamente a média e o desvio padrão da variável $X_i, i = 1, 2, \dots, p$. A matriz $P_{p \times p}$ é a matriz de covariâncias do vetor aleatório $Z = (Z_1 Z_2 \dots Z_p)'$.

A partir da matriz de correlação $P_{p \times p}$ podemos construir o modelo que relaciona linearmente as variáveis padronizadas e os m fatores comuns desconhecidos, como pode-se observar abaixo:

$$\begin{aligned} Z_1 &= l_{11}F_1 + l_{12}F_2 + \dots + l_{1m}F_m + \varepsilon_1 \\ Z_2 &= l_{21}F_1 + l_{22}F_2 + \dots + l_{2m}F_m + \varepsilon_{12} \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ Z_p &= l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + \varepsilon_{1p} \end{aligned}$$

As variáveis latentes correspondem ao vetor $F_{m \times 1}$ contendo m fatores que descrevem os elementos da amostra e são não observáveis. As variáveis Z_i estão relacionadas linearmente com novas variáveis aleatórias $F_j, j = 1, 2, \dots, m$ que serão identificadas de algum modo. O vetor de erros aleatórios $\varepsilon_{p \times 1}$ é a variação de Z_i que não é explicada pelos fatores comuns F_j . As cargas fatoriais l_{ij} são dados pelo coeficiente da i -ésima variável padronizada Z_i no j -ésimo fator F_j e representa o grau de relacionamento linear entre Z_i e F_j . Com isso o objetivo é identificar as novas m -variáveis, interpretá-las e calcular seus escores assim como foi feito na análise de componentes principais.

Rotação Ortogonal de Fatores

Na tentativa de se alcançar uma estrutura mais simples de ser interpretada pode-se utilizar a rotação ortogonal dos fatores. Em alguns casos a dificuldade de interpretabilidade dos fatores é gerada pelo fato dos coeficientes $\hat{l}_{ij}$ apresentarem grandezas numéricas similares, ou seja, a suposição de ortogonalidade dos fatores está sendo violada e a participação das variáveis originais em m grupos não é clara ou difícil de ser justificada.

Segundo Thurstone (1947) a interpretação dos fatores seria possível se uma transformação fizesse com que as cargas fatoriais de cada característica de $Z_i, i = 1, 2, \dots, p$ tivessem um valor numérico grande em somente um dos fatores e valores pequenos ou moderados nos fatores restantes. Com isso as variáveis originais Z_i ficariam divididas em grupos, onde cada grupo estaria correlacionado com apenas um fator, logo cada variável apresentará cargas fatoriais próximos de zero para um grande número de

fatores, e a informação de dois fatores apresentará um padrão diferenciado de cargas fatoriais altas e baixas.

A rotação ortogonal traz benefícios, mas em termos de qualidade de ajuste não acrescenta em nenhuma melhoria em relação ao ajuste obtido usando a matriz de cargas fatoriais ($\widehat{L_{pxm}}$), pois a matriz residual original não é alterada pela transformação ortogonal. Porém quando a solução sem rotação já é de boa qualidade, não se recomenda o uso de rotação ortogonal pelo fato de que a tendência da solução rotacionada ser pior que a solução sem rotação.

Critério Varimax

O critério Varimax foi criado por Kaiser em 1958 e tem como objetivo determinar a matriz ortogonal $T_{m \times m}$ que tem como alicerce a tentativa de encontrar fatores com grandes variabilidades nas cargas fatoriais, ou seja, determinar um fator fixo, um grupo de variáveis Z_i altamente correlacionadas com o fator e outro grupo em que as variáveis estejam correlacionadas desprezivelmente ou moderadamente com cada fator. O método é obtido com a maximização dos quadrados das cargas fatoriais originais das colunas da matriz de cargas fatoriais $\widehat{L_{pxm}}$.

Após a rotação seja $\hat{l}_{ij}^*$ o coeficiente da i -ésima variável no j -ésimo fator e V a quantidade definida por:

$$V = \frac{1}{p} \sum_{j=1}^m \left[\sum_{i=1}^p \tilde{l}_{ij} - \frac{1}{p} \left(\sum_{i=1}^p \tilde{l}_{ij}^2 \right) \right]^2$$

onde , sendo $\hat{h}_i$ a raiz quadrada da comunalidade da variável $Z_i, i = 1, 2, \dots, p$. A quantidade V é proporcional à soma das variâncias das cargas fatoriais escalonadas $\tilde{l}_{ij}$ ao quadrado para cada fator. Esse escalonamento das cargas fatoriais é feito devido as variáveis Z_i não terem necessariamente comunalidades iguais. Portanto o critério Varimax seleciona os coeficientes $\tilde{l}_{ij}$ que maximizam a quantidade V , obtendo assim coeficientes finais da matriz de cargas fatoriais ($\hat{L}^* = \hat{L}T$).

Alfa de Cronbach

O coeficiente alfa de Cronbach foi criado por Lee J. Cronbach (Cronbach, 1951) e tem como objetivo medir a confiabilidade de um instrumento e o grau de covariância dos itens entre si. Refere-se ao grau de inter-correlação entre os itens em uma escala e avalia a homogeneidade dos itens que constituem cada uma das escalas multi-itens. Sendo assim

coeficientes maiores que 0,7 são considerados satisfatórios para estudos de grupo (CRAMER, ILAE REPORT, 2007).

Para Leontitsis e Pagge (2007) a estimação do alfa de Cronbach é feita considerando X uma matriz $m \times k$ que corresponde as respostas quantificadas de um questionário, sendo cada linha um sujeito e cada coluna uma questão. A equação que expressa esse coeficiente é:

$$\alpha = \frac{K}{K-1} \left[\frac{\sigma_{\tau}^2 - \sum_{i=1}^K \sigma_i^2}{\sigma_{\tau}^2} \right]$$

onde σ_i^2 é a variância de cada coluna de X , isto é, é a variância relacionada a cada questão de X , e σ_{τ}^2 é a variância de cada linha de X , ou seja, é a variância da soma de cada resposta de cada sujeito. Uma observação importante é que k deve ser maior do que 1 para que não haja zero no denominador e n deve ser maior do que 1 para que não haja zero no denominador no cálculo do σ_{τ}^2 e do σ_i^2 .

3. Aplicações

Foram convidados a participar do estudo 1000 pessoas portadores de doenças crônicas com idade maior ou igual a 21 anos que estiveram em acompanhamento no ambulatório referente a cada uma delas no Hospital das Clínicas da Faculdade de Medicina da Universidade Federal de Uberlândia (HCU), conforme parecer: 952/10. Esses dados fazem parte do projeto de dissertação de mestrado do aluno Bruno Simão Teixeira e os demais membros do projeto são: Tânia Maria da Silva Mendonça, Carlos Henrique Martins da Silva e Rogério de Melo Costa Pinto.

Trata-se de um estudo transversal, ou seja, os pacientes foram alocados em grupos específicos a cada doença e responderam os itens traduzidos e adaptados culturalmente do PROMIS que foram relevantes para sua doença e um questionário estruturado de dados sociodemográficos. As avaliações médicas de rotina permaneceram no ambulatório específico a cada uma das doenças crônicas no HCU. A aplicação dos instrumentos é realizada nesses mesmos locais.

Participaram da pesquisa pacientes com doenças crônicas diagnosticadas segundo critérios específicos a cada um com idade igual ou superior a 21 anos e que assine o Termo de Consentimento Livre e Esclarecido concordando em participar do estudo. Pacientes com idade igual ou superior a 60 anos e que apresentaram déficit cognitivo segundo o Mini Exame do Estado Mental (MEEM) foram excluídos da pesquisa. Os participantes

aptos para a pesquisa responderam os questionários sociodemográficos e o banco de itens Fadiga.

Os 82 itens são divididos em duas dimensões: 1) 47 itens que avaliam o impacto da fadiga (em atividades físicas, mentais e sociais) os quais avaliam o quanto a fadiga interferiu na sua vida cotidiana; e 2) 35 itens que avaliam a experiência com a fadiga em relação à sua intensidade, frequência e duração. Escalas de Likert com cinco classificações são utilizadas para avaliar a intensidade (Nem um pouco / Um pouco / Mais ou menos / Muito / MUITÍSSIMO); a frequência (Nunca / Raramente / Algumas vezes / Muitas vezes / Sempre); e a duração (Nenhum / 1 dia / 2 – 3 dias / 4 – 5 dias / 6 – 7 dias) (PROMIS®CooperativeGroup, 2008).

Foram observados 1000 questionários sociodemográficos. O sexo feminino representou a maioria 70,8% e o masculino 29,2% dos pacientes, compreendendo uma idade média de 38 anos (desvio = 12,624, mínimo de 18 máximo de 78 anos). Os questionados apresentaram em sua pluralidade serem casados (49,8%), com nível educacional médio/superior completo (70%). Essas informações podem ser observadas na Tabela 1.

Tabela 1: Distribuição da frequência (n) e porcentagem (%) dos dados sociodemográficos de 200 pacientes.

Variáveis	N	(%)
Sexo		
Masculino	292	(29,2)
Feminino	707	(70,8)
Idade média (anos)	38	
Estado Civil		
Solteiro	352	(35,2)
Casado	498	(49,8)
Separado	58	(5,8)
Viúvo	35	(3,5)
Amasiado	56	(5,6)
Nível Educacional		
Fundamental/Completo	300	(30,0)
Médio/Superior	700	(70,0)
Completo		

Avaliando a questão étnico-racial, constatou-se que 51,9% dos indivíduos consideram-se brancos, 33,3% consideram-se pardos, 10,7% consideram-se negros e os 4,1% restantes consideram-se amarelos. A tabela 2 sumariza esses resultados.

Tabela 2: Distribuição da frequência (n) e porcentagens (%) dos dados étnico-raciais.

Variável	N	(%)
Cor		
Amarela	41	(4,1)
Branca	519	(51,9)
Negra	107	(10,7)
Parda	333	(33,3)

Com relação ao vínculo empregatício e a aposentadoria dos pacientes, observamos que 50,8% estão no mercado de trabalho e 16,3% estão aposentados. As informações são perceptíveis na Tabela 3.

Analisando a religião e o envolvimento dos entrevistados, obteve-se que 58% são católicos, 24,8% evangélicos, 10,4% espíritas, 6,1% declaram-se sem religião, 0,6% ocultaram-se sua religião e 0,1% são budistas. Com relação ao envolvimento na religião, 39,3% diz-se ter um “moderado” comprometimento, 36,5% tem um “alto” encantamento, 17,3% pouco se relaciona e 6,9% não possui envolvimento. Essas informações podem ser observadas na Tabela 4.

Tabela 3: Distribuição da frequência (n) e porcentagens (%) dos dados da aposentadoria e do vínculo empregatício.

Variáveis	N	(%)
Empregado		
Não	492	(49,2)
Sim	508	(50,8)
Aposentado		
Não	837	(83,7)
Sim	163	(16,3)

Tabela 4: Distribuição da frequência (n) e porcentagens (%) dos dados da religião.

Variáveis	N	(%)
Religião		
Católica	580	(58,0)
Espírita	104	(10,4)
Evangélica	248	(24,8)
Nenhuma	61	(6,1)
Ocultismo	6	(0,6)
Budismo	1	(0,1)
Envolvimento		
Alto	365	(36,5)
Baixo	173	(17,3)
Moderado	393	(39,3)
Nenhum	69	(6,9)

A maioria dos pacientes (60%) se declaram saudáveis, sendo apenas 40% doentes que em média não apresentam doenças crônicas. Apenas 38,5% tomam medicamentos diariamente, o que é possível ser analisado na Tabela 5.

Tabela 5: Distribuição da frequência (n) e porcentagem dos dados dos pacientes Saudáveis/Doentes.

Variáveis	N	(%)
Doente	400	(40,0)
Saudável	600	(60,0)
Número de Doenças		
0	620	(62,0)
1	293	(29,3)
2	72	(7,2)
3	13	(1,3)
4	2	(0,2)
Medicação		
Não	615	(61,5)
Sim	385	(38,5)

A Tabela 6 refere-se as questões 1 e 2 do WHOQOL-BREF, onde foi analisado a satisfação geral sobre a qualidade de vida dos pacientes e sua saúde. Foi observado positivamente na Questão 1 que a maioria dos indivíduos possui uma qualidade de vida boa (53%). Na Questão 2 mostra a satisfação dos pacientes quanto a sua saúde, sendo assim a maioria satisfeito (45,6%).

Tabela 6: Distribuição da frequência (n) e porcentagem (%) dos dados das Questões 1 e 2 do WHOQOL-BREF.

Variáveis	N	(%)
Questão 1		
Muito Ruim	9	(0,9)
Ruim	42	(4,2)
Nem Ruim/Nem Boa	206	(20,6)
Boa	530	(53,0)
Muito Boa	213	(21,3)
Questão 2		
Muito Insatisfeito	26	(2,6)
Insatisfeito	86	(8,6)
NemSatisfeito/Nem	219	(21,9)
Insatisfeito		
Satisfeito	456	(45,6)
Muito Satisfeito	213	(21,3)

Com base nos resultados obtidos pela técnica dos componentes principais, os respectivos autovalores e porcentagens da variância explicada por cada um estão apresentados na Tabela 7. Os três primeiros componentes principais foram responsáveis por 57,581% da variação total, sobre a fadiga dos pacientes, em que o componente um (CP_1) foi responsável por 50,563%, o segundo, (CP_2), por 3,593% das variações dos dados e o terceiro (CP_3) por 3,425%.

Verificou-se que como os três primeiros componentes principais gerados a partir desta análise que tem autovalores > 1 ($\lambda_i > 1$) (Kaiser, 1958; FRAGA, et al., 2015) e foram responsáveis por 57,581% da variância total no conjunto de dados os três componentes principais foram retidos, com o auxílio do screeplot (Figura 2) e estão apresentados na Tabela 7. Portanto, os três primeiros componentes principais resumem efetivamente a variância amostral total e podem ser utilizados para o estudo do conjunto de dados.

Tabela 7: Componentes principais (CPs), autovalores (λ_i) e porcentagem da variância explicada e proporção acumulada (%) pelos componentes.

Componente Principal	Autovalores	Proporção (%)	Proporção Acumulada (%)
CP_1	41,462	50,563	50,563
CP_2	2,946	3,593	54,156
CP_3	2,809	3,425	57,581
CP_4	1,476	1,800	59,381
CP_5	1,225	1,494	60,874
CP_6	1,083	1,321	62,195
CP_7	1,034	1,261	63,456
CP_8	0,967	1,180	64,636
CP_9	0,964	1,175	65,811
CP_{10}	0,884	1,078	66,888

Figura 2: O *screeplot* dos autovalores dos componentes principais.

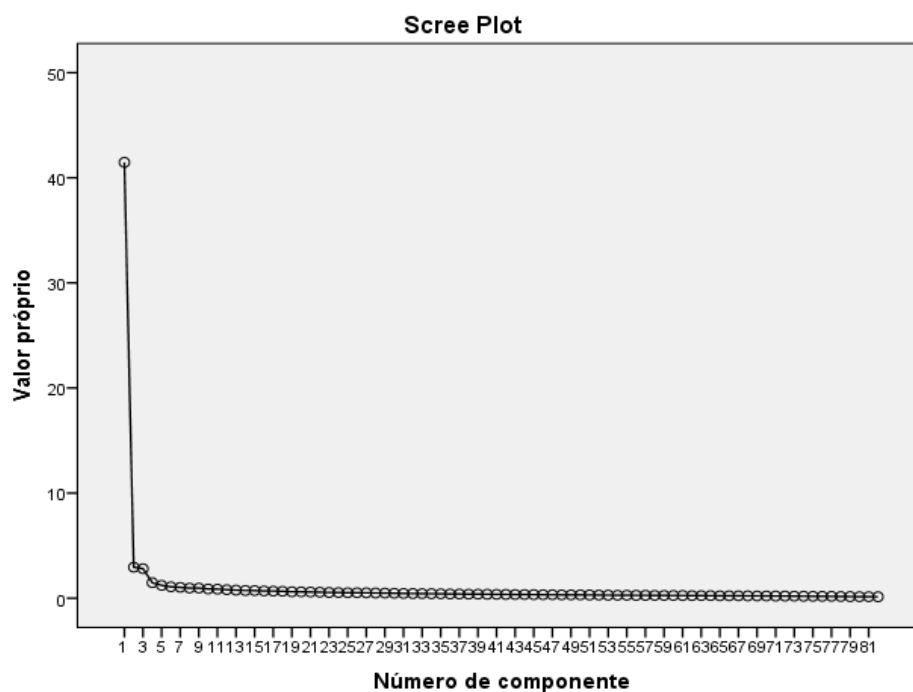


Tabela 8: Coeficientes de ponderação das características com os três primeiros componentes principais.

Variável	CP1	CP2	CP3	Variável	CP1	CP2	CP3	Variável	CP1	CP2	CP3
FAD01	0,664			FAD31	0,792			FAD61	0,715		
FAD02	0,667			FAD32	0,7642			FAD62	0,551		
FAD03	0,723			FAD33	0,7656			FAD63	0,672		
FAD04	0,726			FAD34		0,301	0,439	FAD64		0,4133	0,5828
FAD05	0,745			FAD35	0,7831			FAD65	0,699		
FAD06	0,743			FAD36	0,7985			FAD66	0,795		
FAD07	0,591			FAD37	0,6312			FAD67	0,766		
FAD08	0,705			FAD38	0,7856			FAD68	0,755		
FAD09	0,745			FAD39	0,766			FAD69	0,799		
FAD10	0,747			FAD40	0,7942			FAD70	0,746		
FAD11	0,568			FAD41	0,7976			FAD71		0,3887	0,6086
FAD12	0,777			FAD42	0,8218			FAD72	0,654		
FAD13	0,746			FAD43	0,826			FAD73		0,3825	0,6259
FAD14	0,751			FAD44	0,7827			FAD74	0,731		
FAD15	0,73			FAD45	0,588			FAD75	0,664		
FAD16	0,771			FAD46	0,8025			FAD76	0,754		
FAD17	0,503			FAD47	0,7833			FAD77	0,681		
FAD18	0,761			FAD48	0,785			FAD78	0,626		
FAD19	0,582			FAD49	0,7349			FAD79	0,692		
FAD20	0,781			FAD50	0,7448			FAD80	0,636		
FAD21	0,77			FAD51	0,7626			FAD81		0,4278	0,5233
FAD22	0,708			FAD52	0,7419			FAD82	0,786		
FAD23	0,673			FAD53	0,7382						
FAD24	0,794			FAD54	0,7667						
FAD25	0,67			FAD55	0,7678						
FAD26	0,783			FAD56	0,7866						
FAD27	0,773			FAD57	0,7627						
FAD28	0,789			FAD58	0,7869						
FAD29	0,812			FAD59	0,7486						
FAD30	0,812			FAD60			0,537				

Tabela 9: Coeficientes de correlações dos três primeiros componentes principais

Variável	Correlação	Variável	Correlação	Variável	Correlação
FAD01	0,44781754	FAD31	0,66035414	FAD61	0,538124781
FAD02	0,47925087	FAD32	0,598168155	FAD62	0,326312442
FAD03	0,55759982	FAD33	0,60993482	FAD63	0,547706254
FAD04	0,57631195	FAD34	0,283288449	FAD64	0,587751342
FAD05	0,58818447	FAD35	0,629868913	FAD65	0,616993539
FAD06	0,63327548	FAD36	0,66390385	FAD66	0,68103083
FAD07	0,37369504	FAD37	0,413189341	FAD67	0,722048789
FAD08	0,60339154	FAD38	0,623985673	FAD68	0,644689487
FAD09	0,61958793	FAD39	0,588004625	FAD69	0,707862624
FAD10	0,62669978	FAD40	0,64089051	FAD70	0,657578248
FAD11	0,33481867	FAD41	0,636420443	FAD71	0,536427062
FAD12	0,6635909	FAD42	0,678969091	FAD72	0,548612037
FAD13	0,58460076	FAD43	0,685458963	FAD73	0,59674493
FAD14	0,60478677	FAD44	0,617628525	FAD74	0,55194705
FAD15	0,61158232	FAD45	0,352779901	FAD75	0,461601196
FAD16	0,6416899	FAD46	0,644404959	FAD76	0,62751657
FAD17	0,27732082	FAD47	0,61454267	FAD77	0,47533264
FAD18	0,61794019	FAD48	0,66608768	FAD78	0,410262085
FAD19	0,37908927	FAD49	0,614672259	FAD79	0,549433132
FAD20	0,6604754	FAD50	0,598106698	FAD80	0,481215635
FAD21	0,66345199	FAD51	0,608001192	FAD81	0,523608663
FAD22	0,52530568	FAD52	0,599109364	FAD82	0,640527176
FAD23	0,47056126	FAD53	0,636638952		
FAD24	0,65874085	FAD54	0,613005467		
FAD25	0,4773479	FAD55	0,627008422		
FAD26	0,62048861	FAD56	0,700092824		
FAD27	0,61694608	FAD57	0,657608849		
FAD28	0,64227411	FAD58	0,657220491		
FAD29	0,67360614	FAD59	0,586611624		
FAD30	0,68460419	FAD60	0,362255817		

Na Tabela 8 e 9, são apresentados, ainda, as correlações de cada variável e seus coeficientes de ponderação de cada característica. Como intuito de se entender a importância de cada variável na construção dos três componentes foi calculado a correlação entre as variáveis originais mostrando o seu grau de importância para o modelo, sendo que a maioria das variáveis apresentaram um grau de correlação moderado, como foram apresentados na Tabela 9.

Através do método de item deletado do Alfa de Cronbach decidiu-se em optar em deixar as variáveis $X_{34}, X_{64}, X_{71}, X_{73}, X_{81}$ relacionadas ao componente 2 devido terem apresentado uma maior confiabilidade no valor de 0,768 nos mostrando uma maior correlação com o componente, ou seja, uma maior inter-correlação entre os itens.

Com a seleção de três componentes principais, a redução da dimensão de 82 variáveis originais para três componentes principais é bastante razoável. Portanto decidiu-se utilizar unicamente os três primeiros componentes principais para a composição das equações 1, 2 e 3.

$$\begin{aligned}
 CP_1 = & 0,664X_1 + 0,667X_2 + 0,723X_3 + 0,726X_4 + 0,745X_5 + 0,743X_6 + 0,591X_7 + 0,705X_8 \\
 & + 0,745X_9 + 0,747X_{10} + 0,568X_{11} + 0,777X_{12} + 0,746X_{13} + 0,751X_{14} + 0,730X_{15} \\
 & + 0,771X_{16} + 0,503X_{17} + 0,761X_{18} + 0,582X_{19} + 0,781X_{20} + 0,770X_{21} + 0,708X_{22} \\
 & + 0,673X_{23} + 0,794X_{24} + 0,670X_{25} + 0,783X_{26} + 0,773X_{27} + 0,789X_{28} + 0,812X_{29} \\
 & + 0,812X_{30} + 0,792X_{31} + 0,764X_{32} + 0,766X_{33} + 0,783X_{35} + 0,799X_{36} + 0,631X_{37} \\
 & + 0,786X_{38} + 0,766X_{39} + 0,794X_{40} + 0,798X_{41} + 0,822X_{42} + 0,826X_{43} + 0,783X_{44} \\
 & + 0,588X_{45} + 0,803X_{46} + 0,783X_{47} + 0,785X_{48} + 0,735X_{49} + 0,745X_{50} + 0,763X_{51} \\
 & + 0,742X_{52} + 0,738X_{53} + 0,767X_{54} + 0,768X_{55} + 0,787X_{56} + 0,763X_{57} + 0,787X_{58} \\
 & + 0,749X_{59} + 0,715X_{61} + 0,551X_{62} + 0,672X_{63} + 0,699X_{65} + 0,795X_{66} + 0,766X_{67} \\
 & + 0,755X_{68} + 0,799X_{69} + 0,746X_{70} + 0,654X_{72} + 0,731X_{74} + 0,664X_{75} + 0,754X_{76} \\
 & + 0,681X_{77} + 0,626X_{78} + 0,692X_{79} + 0,636X_{80} + 0,786X_{82} \quad (1)
 \end{aligned}$$

$$CP_2 = 0,301X_{34} + 0,413X_{64} + 0,389X_{71} + 0,382X_{73} + 0,428X_{81} \quad (2)$$

$$CP_3 = 0,537X_{60} \quad (3)$$

De acordo com a equação (1) e a Tabela 8, no primeiro componente principal destacaram-se todas variáveis apresentando um alto índice de ponderação, sendo assim, observa-se que essas variáveis estão relacionadas a fadiga física e social dos indivíduos e neste caso, o denominamos componente fadiga física e social. E de acordo com a equação (2) e Tabela 8, no segundo componente principal ficou evidente o destaque das variáveis X_{64} (FAD64) e X_{81} (FAD81) que estão relacionados com o grau de energia e a energia física dos indivíduos, podendo assim ser chamado de componente fadiga energia física. Na equação (3) e Tabela 8, observa-se que a variável X_{60} (FAD60) representa o terceiro componente principal e é caracterizada por representar se os indivíduos tiveram energia suficiente para realizar as atividades físicas de lazer, sendo assim pode-se chamá-lo de componente fadiga lazer.

As variáveis do CP_1 apresentaram contribuições similares, isto foi verificado pelas variáveis que têm vetor de maior comprimento e que foram mais próximas ao eixo CP_1 , mostrado na Figura 5. Existem correlações altas entre as variáveis $X_{34}, X_{64}, X_{71}, X_{73}, X_{81}$ e o CP_2 , pois formaram ângulos agudos entre as variáveis, também é perceptível a alta correlação entre CP_1 e as variáveis que são correspondentes a ele, como pode-se observar na Figura 4. Na Figura 3 observa-se uma forte correlação entre as variáveis $X_{34}, X_{64}, X_{71}, X_{73}, X_{81}$ tanto para o CP_2 e o CP_3 , mas também há uma forte correlação da variável X_{60} , justificada pela amplitude dos ângulos que ambas as variáveis apresentaram.

Figura 3: Biplot $CP_2 \times CP_3$ sobre as variáveis Fadiga-Promis pela ACP

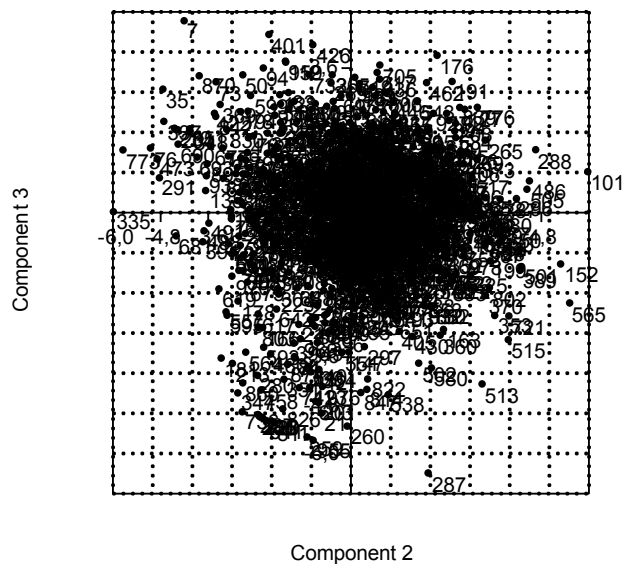


Figura 4: Biplot $CP_1 \times CP_2$ sobre as variáveis Fadiga-Promis pela ACP

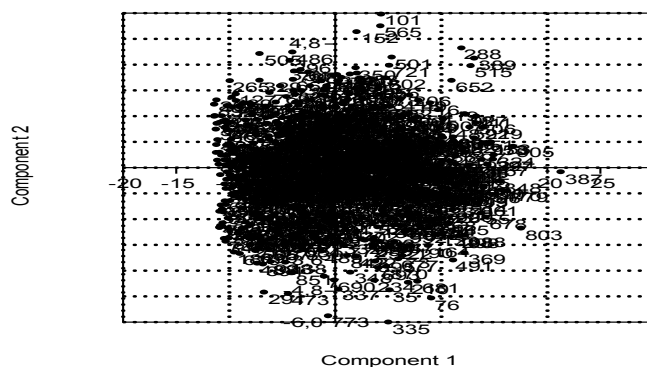
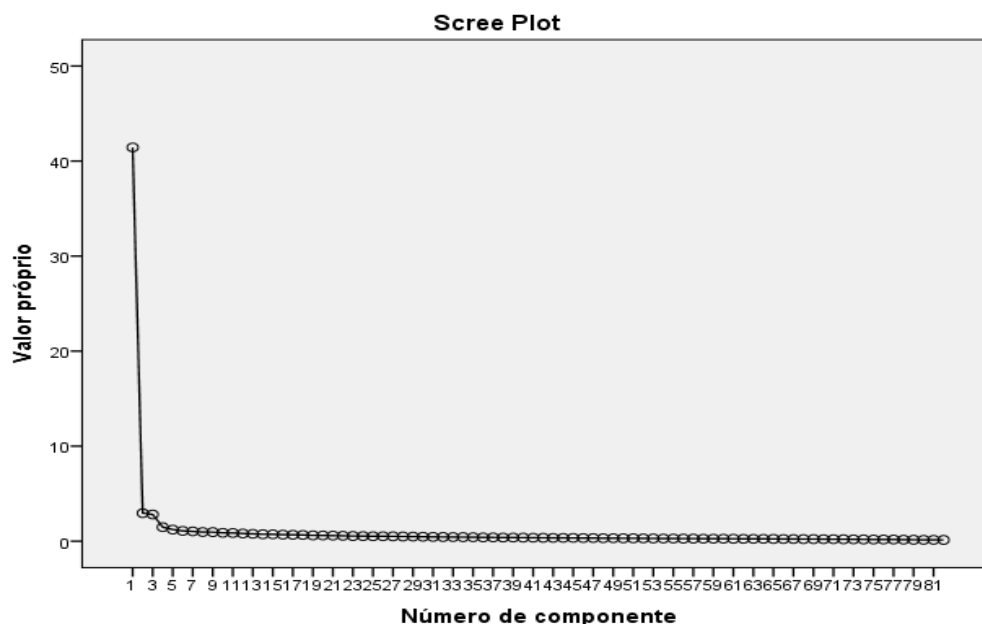


Tabela 10: Análise Fatorial exploratória, autovalores (λ_i) e porcentagem da variância explicada e proporção acumulada (%) pelos componentes.

Fatores	Autovalores	Proporção (%)	Proporção Acumulada (%)
FAT_1	41,432	50,527	50,527
FAT_2	2,953	3,601	54,128
FAT_3	2,814	3,431	57,560
FAT_4	1,478	1,802	59,362
FAT_5	1,227	1,497	60,859
FAT_6	1,085	1,323	62,182
FAT_7	1,035	1,262	63,444
FAT_8	0,969	1,182	64,626
FAT_9	0,959	1,170	65,796
FAT_{10}	0,881	1,074	66,870

Figura 6: O *screeplot* dos autovalores da análise fatorial por meio do método de componentes principais.



Com a dificuldade na interpretabilidade dos fatores originais $F_1, F_2, \dots, F_m$ utilizou-se o recurso da transformação ortogonal dos fatores com o objetivo de que as cargas fatoriais de cada característica de $Z_i, i = 1, 2, \dots, p$ tivessem um valor numérico grande em

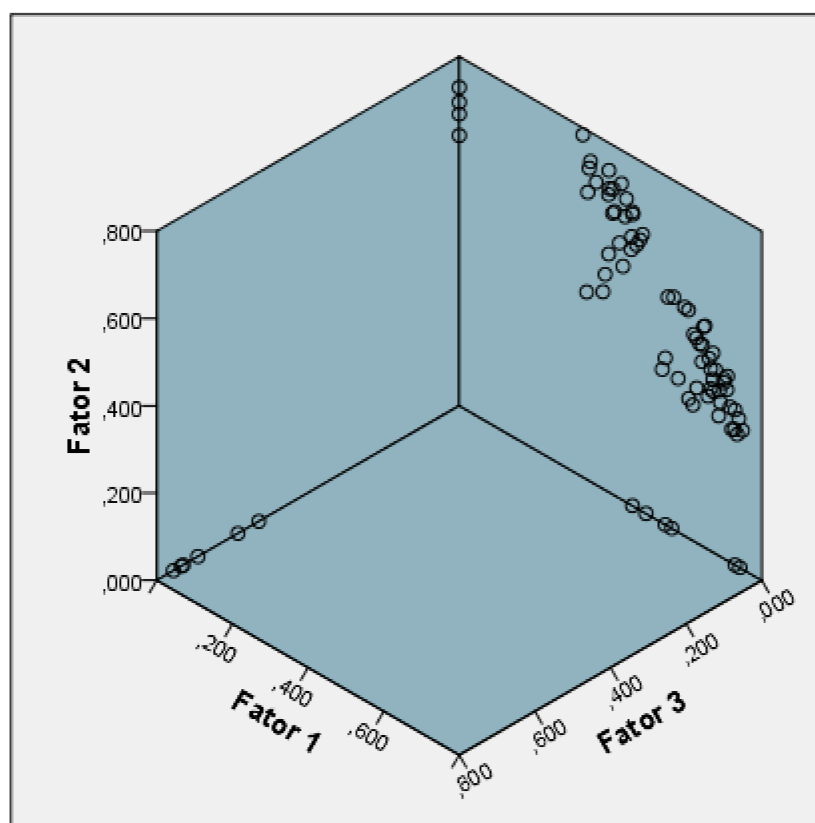
somente um dos fatores e valores pequenos ou moderados nos fatores restantes. Isto só foi possível após utilizar o critério de rotação varimax, que tem como alicerce a tentativa de encontrar fatores com grandes variabilidades nas cargas fatoriais, ou seja, determinar um fator fixo, um grupo de variáveis Z_i altamente correlacionadas com o fator e outro grupo em que as variáveis estejam correlacionadas desprezivelmente ou moderadamente com cada fator. A tabela 11 apresenta os resultados da rotação ortogonal varimax para os três fatores.

Tabela 11: Análise Fatorial Rotacionada através do método varimax, e as respectivas cargas fatoriais de cada variável.

Variável	FAT1	FAT2	FAT3	Variável	FAT1	FAT2	FAT3	Variável	FAT1	FAT2	FAT3
FAD01	0,545	0,382		FAD31	0,7075	0,391		FAD61	0,4243	0,585	
FAD02	0,618	0,311		FAD32	0,6416	0,422		FAD62	0,337	0,429	
FAD03	0,658	0,351		FAD33	0,6706	0,396		FAD63		0,668	
FAD04	0,686	0,32		FAD34			0,53	FAD64			0,735
FAD05	0,672	0,369		FAD35	0,6608	0,439		FAD65		0,729	
FAD06	0,735	0,302		FAD36	0,6995	0,411		FAD66	0,4418	0,695	
FAD07	0,545			FAD37	0,5365	0,352		FAD67	0,3265	0,784	
FAD08	0,742			FAD38	0,6201	0,474		FAD68	0,3973	0,696	
FAD09	0,721	0,307		FAD39	0,5516	0,500		FAD69	0,4294	0,723	
FAD10	0,728	0,31		FAD40	0,6426	0,461		FAD70	0,3473	0,733	
FAD11	0,494			FAD41	0,5963	0,524		FAD71			0,73
FAD12	0,739	0,34		FAD42	0,6459	0,504		FAD72		0,695	
FAD13	0,665	0,372		FAD43	0,6498	0,508		FAD73			0,756
FAD14	0,691	0,353		FAD44	0,6266	0,469		FAD74	0,455	0,585	
FAD15	0,729			FAD45	0,3796	0,451		FAD75	0,3953	0,545	
FAD16	0,716	0,355		FAD46	0,606	0,522		FAD76	0,395	0,681	
FAD17	0,458			FAD47	0,567	0,532		FAD77	0,4327	0,536	
FAD18	0,688	0,379		FAD48	0,4585	0,673		FAD78	0,3859	0,494	
FAD19	0,562			FAD49	0,362	0,693		FAD79	0,3404	0,659	
FAD20	0,729	0,355		FAD50	0,4113	0,65		FAD80		0,619	
FAD21	0,749	0,318		FAD51	0,4688	0,603		FAD81			0,691
FAD22	0,628	0,355		FAD52	0,4055	0,644		FAD82	0,4846	0,635	
FAD23	0,579	0,352		FAD53	0,3443	0,716					
FAD24	0,7	0,403		FAD54	0,478	0,619					
FAD25	0,607	0,32		FAD55	0,4386	0,653					
FAD26	0,636	0,46		FAD56	0,3959	0,737					
FAD27	0,665	0,417		FAD57	0,407	0,7					
FAD28	0,678	0,421		FAD58	0,459	0,668					
FAD29	0,671	0,456		FAD59	0,4559	0,615					
FAD30	0,709	0,422		FAD60			0,585				

Como pode-se ser observado a solução com três fatores relacionados, produz uma melhor distribuição do agrupamento das 82 variáveis. O FAT_1 está correlacionado com as variáveis em negrito na sua respectiva coluna como podemos observar na Tabela 11, que se refere às variáveis de impacto que os indivíduos apresentam sobre a fadiga, sendo assim foi denominado fator fadiga impacto. De acordo com a Tabela 11, o fator FAT_2 está relacionado com o grau de “experiência” dos indivíduos, podendo assim ser chamado de fator fadiga experiência. Na Tabela 11, também observa-se que as variáveis em negrito da coluna FAT_3 que representam o terceiro fator, são caracterizadas por representar os indivíduos que tiveram energia suficiente para realizar as atividades físicas. Assim, o chamamos de componente fadiga energia. Portanto, pode-se confirmar também através da Figura 7 as variáveis que estão inseridas em cada fator.

Figura 7: Gráfico das cargas fatoriais resultantes da transformação varimax.



Aplicou-se o coeficiente de alfa Cronbach no banco de itens Fadiga-Promis com o objetivo de medir a inter-correlação dos itens. Sendo assim foi verificado se os mesmos

estão na mesma dimensão ou no mesmo construto. Obteve-se um coeficiente de 0,984 o que nos indica uma excelente confiabilidade do questionário, indicando que as variáveis estão no mesmo construto, ou seja, elas estão inter-correlacionadas, como pode-se verificar na Tabela 12.

Tabela 12: Consistência do PROMIS pelo alfa de Cronbach.

Questionário	A
Fadiga-Promis	0,984

4. Conclusão

Tendo em vista os resultados obtidos, a análise de componentes principais e a análise fatorial através do método de componentes principais se mostrou efetiva e permitiu a retirada ou descarte de setenta e nove variáveis que apresentaram baixa variabilidade ou foram redundantes por estarem correlacionadas com as de maior importância para três componentes principais e três fatores. Assim, um menor número de variáveis foram necessárias para explicar a variação total resultando em economia de tempo e de recursos em futuros trabalhos que utilizarão essa mesma base de dados, sem grande perda de informação.

Um dos objetivos da análise de componentes principais e a análise fatorial através do método de componentes principais, neste caso, foi atingido, pois um número relativamente pequeno de componentes (CP_1, CP_2 e CP_3) e fatores (FAT_1, FAT_2 e FAT_3) foi extraído com a capacidade de explicar a maior variabilidade nos dados originais (57,581% e 57,56%). Através dos fatores obtidos e da análise do questionário, pudemos dar uma interpretação para os fatores FAT_1, FAT_2 e FAT_3 , a saber, fadiga impacto, fadiga experiência e fadiga energia, respectivamente.

5. Referências

ARTES, R. Aspectos estatísticos da análise fatorial de escalas de avaliação. **Revista de psiquiatria clínica**, São Paulo, v.25, n. 5, p.223-228, 1998.

CATTELL, R. B. (1966). The scree test for the number of factors. **Multivariate Behavioral Research**, 1, 245-276.

CELLA D, YOUNT S, ROTHROCK N, et al. The patient-reported outcome measurement information system (PROMIS): progress of an NIH roadmap cooperative group during its first two years. **Medical Care**. v. 45, (Suppl 1), p. 3-11, 2007.

CICONELLI, R. M. et al. Tradução para a língua portuguesa e validação do questionário genérico de avaliação de qualidade de vida SF-36 (Brasil SF-36). **Revista Brasileira de Reumatologia**, São Paulo, v. 39, n. 3, p. 143-150, mai/jun. 1999.

CRONBACH, L. J. Coefficient alpha and the internal structure of test. **Psychometrika**. 1951.

EISER, C.; MORSE, R. Quality-of-life measures in chronic diseases of childhood. **Health Technology Assessment**, v. 5, n. 4, p. 10-15, 2001.

FRIES, J.F.; BRUCE, B.; CELLA, D. The promise of PROMIS: Using item response theory to improve assessment of patient-reported outcomes. **Clin. Exp Rheumatol**. v. 23 (suppl. 39), p. 53-7, 2005.

GABRIEL, K.R. The biplot graphic display of matrices with application to principal componente analysis. **Biometrika**, v.58, n.3, p.453-467, 1971.

GRIEP, R. H. et al. Confiabilidade teste-reteste de aspectos da rede social no Estudo Pró-Saúde. **Revista Saúde Pública**, São Paulo, v. 37, n.3, p. 379-85, 2003.

GUYATT, G. H.; FEENY, D. H.; PATRICK, D. L. Measuring Health-related Quality of Life. **Annals of Internal Medicine**, Philadelphia, v. 118, n. 8, p. 622-629, Apr. 1993.

HAIR, J. F. JR. et al, Análise multivariada de dados. 5. ed, Porto Alegre: **Bookman**, 2005. 593p.

HARDLE, W.; SIMAR, L. Applied Multivariate Statistical Analysis. **Local: Tech**, 2003. p.275-298.

KRUPP, L.B.; LAROCCA, N.C.; MUIR-NASH, J. The fatigue severity scale: application to patients with multiples esclerosis and systemic lupus erythematosus. **Archive Neurologic**. v. 46, p.1121-3, 1989.

LEONTITSIS, A.; PAGGE, J. A simulation approach on Cronbach's alpha statistical significance. **Mathematics and Computers in Simulation**. v. 73, p. 336-340. 2007.

McCARTHY, M. L. et al. Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. **Developmental Medicine and Child Neurology**, London, v. 44, n. 7, p. 468-476, July 2002.

McHORNEY, C. A. et al. The MOS 36-item short form health survey (SF-36): III. Test of data quality, scaling assumptions, and reliability across diverse patient groups. **Medical Care**, Philadelphia, v. 32, n. 1, p. 40-66, Jan. 1994.

MOTA, N.F.; CRUZ, D.A.L.M.; FINI, A. Fadiga em adultos acompanhantes de pacientes em tratamento ambulatorial. **Acta Paulista de Enfermagem**. v. 23, n. 3, p. 348-53, 2010.

PEÑA, E.D. Lost in Translation: Methodological considerations in cross-cultural research. **Child Development**. v. 78, n. 4, p. 1255-64, 2007.

PROMIS Cooperative Group. **Unpublished Manual for the Patient-Reported Outcomes Measurement Information System (PROMIS) Version 1.1**, October, 2008.

RENCHE, A. C. **Methods of Multivariate Analysis. Factor Analysis**, 2.ed. : Wiley-Interscience, 2001. p.408-448.

STRAPASSON, E.; VENCovsky, R.; BATISTA, L. A. R. **Seleção de Descritores na Caracterização de Germoplasma de Paspalum sp. por meio de Componentes Principais**. Rev. bras. zootec., 29(2):373-381, 2000.

THURSTONE. (1947), **Multiple Factor Analysis**, Univ. of Chicago Press 1947.

WARE J.E. JR; KOSINSKI, M.; BJORNER J.B. Item banking and the improvement of health status measures. **Qual Life Res.**v. 2, p.2–5. 2004.

YORKSTON, K.M.; JOHNSON, K.; BOESFLUG, E. SKALA, J.; AMTMANN, D. Communicating about the experience of pain and fatigue in disability. **Quality of Life Research**. v. 19, p. 243–51, 2010

6.Apêndice

QUESTIONÁRIOS

DADOS SÓCIODEMOGRÁFICOS

- 1) Idade: _____
- 2) Sexo: () F () M
- 3) Escolaridade: _____
- 4) Estado civil:
() Casado () Solteiro () Viúvo () Separado () Amasiado
- 5) Raça:
() Branca () Negra () Parda () Amarela () Indígena
- 6) Município que reside: _____
- 7) Está empregado no momento?
() Sim. Atividade profissional que trabalha _____ () Não
- 8) Recebe aposentadoria?
() Sim. Motivo da aposentadoria _____ () Não
- 9) Renda familiar em reais: _____
- 10) Religião a que pertence:
() Católica () Afro-brasileira
() Evangélica () Judaísmo
() Espírita () Outras _____
() Budismo () Nenhuma
() Ortodoxa

11) Grau de envolvimento com a religião:

- ☐ Nenhum
- ☐ Baixo
- ☐ Moderado
- ☐ Alto

12) Possui doença crônica? ☐ Sim ☐ Não

- ☐ Hipertensão arterial ☐ Diabetes ☐ Depressão ☐ Asma ☐ DPOC
- ☐ Doenças neurológicas ☐ Doenças auto-imunes ☐ Doenças reumatológicas

13) Possui outras doenças? ☐ Sim ☐ Não

- ☐ Chagas ☐ Coluna ☐ Cardiopatias ☐ Anemia falciforme
- ☐ Osteoporose

14) Faz o uso de alguma medicação:

- ☐ Sim Qual(is)? _____ ☐ Não

BANCO DE ITENS FADIGA DO PROMIS® -**PROMIS® Fadiga – versão brasileira****Por favor, responda a cada item marcando apenas uma resposta por item.****Nos últimos 7 dias...**

01 FADIMP1	Até que ponto você teve que se esforçar para conseguir fazer as coisas por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
02 FADIMP2	Até que ponto o seu raciocínio ficou mais lento por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
03 FADIMP3	Com que frequência você teve que se esforçar para conseguir fazer as coisas por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
04 FADIMP4	Com que frequência a fadiga interferiu nas suas atividades sociais?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
05 FADIMP5	Com que frequência você ficou menos eficiente no trabalho por causa da fadiga (incluindo o trabalho em casa)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
06 FADIMP6	Com que frequência o seu raciocínio ficou mais lento por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
07 FADIMP8	Com que frequência você ficou cansado(a) demais para ver televisão?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
08 FADIMP9	Com que frequência a fadiga fez com que você tivesse dificuldade em planejar (planejar) atividades com antecedência?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
09 FADIMP10	Com que frequência você teve dificuldade em começar algo novo por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
10 FADIMP11	Com que frequência a fadiga fez com que você ficasse mais esquecido(a)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
11 FADIMP13	Com que frequência você ficou cansado(a) demais para fazer tarefas fora de casa?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
12 FADIMP14	Com que frequência a fadiga dificultou a organização dos seus pensamentos ao fazer as coisas no trabalho (incluindo o trabalho em casa)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
13 FADIMP15	Com que frequência a fadiga interferiu na sua capacidade de participar em atividades de lazer?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
14 FADIMP16	Com que frequência você teve dificuldade em terminar coisas por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
15 FADIMP17	Com que frequência teve dificuldade em tomar decisões por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
16	Com que frequência você teve que limitar as suas atividades sociais por causa da fadiga?	Nunca	Raramente	Algumas	Muitas	Sempre

FADIMP18				vezes	vezes	
17 FADIMP19	Com que frequência você ficou cansado(a) demais para fazer tarefas domésticas?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
18 FADIMP20	Com que frequência você ficou menos alerta por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
19 FADIMP21	Com que frequência você esteve cansado(a) demais para tomar um banho de banheira ou de chuveiro?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
20 FADIMP22	Com que frequência a fadiga dificultou a organização dos seus pensamentos ao fazer as coisas em casa?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
21 FADIMP24	Com que frequência você teve dificuldade em começar coisas por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
22 FADIMP25	Com que frequência foi um esforço manter uma conversa por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
23 FADIMP26	Com que frequência você ficou cansado(a) demais para socializar com a família?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
24 FADIMP27	Até que ponto você teve dificuldade em começar coisas por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
25 FADIMP28	Até que ponto você teve dificuldade em manter uma conversa por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
26 FADIMP29	Com que frequência você ficou cansado(a) demais para sair de casa?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
27 FADIMP30	Com que frequência você ficou cansado(a) demais para pensar com clareza?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
28 FADIMP33	Com que frequência a fadiga o/a limitou no trabalho (incluindo trabalho em casa)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
29 FADIMP34	Até que ponto você teve que limitar as suas atividades sociais por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
30 FADIMP35	Até que ponto você teve dificuldade para organizar os seus pensamentos ao fazer as coisas em casa por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
31 FADIMP36	Até que ponto você teve dificuldade para começar algo novo por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
32 FADIMP37	Você foi menos eficiente no trabalho (incluindo trabalho em casa) por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
33 FADIMP38	Até que ponto você teve dificuldade em tomar decisões por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
34	Com que frequência você teve energia suficiente para se exercitar com vigor?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre

FADIMP40				vezes	vezes	
35 FADIMP42	Com que frequência você foi menos eficiente em casa por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
36 FADIMP43	Até que ponto a fadiga dificultou a organização dos seus pensamentos ao fazer as coisas no trabalho (incluindo trabalho em casa)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
37 FADIMP44	Até que ponto a fadiga fez com que você ficasse mais esquecido(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
38 FADIMP45	Até que ponto a fadiga interferiu na sua capacidade de participar de atividades de lazer?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
39 FADIMP47	Até que ponto você teve que se esforçar para se levantar da cama e fazer as coisas por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
40 FADIMP48	Até que ponto a fadiga interferiu nas suas atividades sociais?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
41 FADIMP49	Até que ponto a fadiga interferiu no funcionamento do seu corpo?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
42 FADIMP50	A fadiga fez com que ficasse menos eficiente em casa?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
43 FADIMP51	Até que ponto você teve dificuldade em terminar coisas por causa da fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
44 FADIMP52	Até que ponto a fadiga fez com que você ficasse menos alerta?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
45 FADIMP53	Com que frequência você ficou cansado(a) demais para fazer uma caminhada curta?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
46 FADIMP55	Com que frequência você teve que se esforçar para se levantar e fazer as coisas por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
47 FADIMP56	Com que frequência você esteve cansado demais para socializar com amigos?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
48 FADEXP2	Com que frequência você se sentiu desgastado(a)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
49 FADEXP5	Com que frequência você sentiu exaustão extrema?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
50 FADEXP6	Com que frequência você se sentiu cansado(a) mesmo não tendo feito nada?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
51 FADEXP7	Com que frequência você sentiu que a fadiga estava fora de controle?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
52	Até que ponto você se sentiu cansado(a) mesmo não tendo feito nada?	Nem um	Um pouco	Mais ou	Muito	Muitíssimo

FADEXP12		pouco		menos		
53 FADEXP13	Em média, até que ponto você ficou física e emocionalmente esgotado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
54 FADEXP16	Com que frequência você se sentiu lento(a)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
55 FADEXP18	Com que frequência você ficou sem energia?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
56 FADEXP19	Com que frequência você ficou fisicamente esgotado(a)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
57 FADEXP20	Com que frequência você se sentiu cansado(a)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
58 FADEXP21	Até que ponto você ficou fatigado(a) quando a fadiga estava no seu pior?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
59 FADEXP22	Com que frequência você se sentiu incomodado(a) por causa da fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
60 FADEXP24	Com que frequência você teve energia suficiente para apreciar as atividades de lazer?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
61 FADEXP26	Com que frequência você ficou cansado(a) demais para aproveitar a vida?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
62 FADEXP28	Com que frequência você ficou cansado(a) demais para se sentir feliz?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
63 FADEXP29	Com que frequência você se sentiu totalmente esgotado(a)?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
64 FADEXP31	Com que frequência você ficou cheio(a) de energia?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
65 FADEXP34	Em média, até que ponto você se sentiu cansado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
66 FADEXP35	Em média, até que ponto você se sentiu incomodado(a) pela fadiga?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
67 FADEXP36	Em média, até que ponto você ficou exausto(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
68 FADEXP38	Até que ponto você ficou fatigado(a) no dia em que se sentiu mais fatigado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
69 FADEXP40	Em média, até que ponto você se sentiu fatigado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
70	Em média, até que ponto você se sentiu desgastado(a)?	Nem um	Um pouco	Mais ou	Muito	Muitíssimo

FATEXP41		pouco		menos		
71 FATEXP42	Em média, quanta energia mental você teve?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
72 FATEXP43	Em média, até que ponto você ficou fisicamente esgotado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
73 FATEXP44	Em média, até que ponto você ficou cheio(a) de energia?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
74 FATEXP45	Em média, até que ponto você se sentiu lento(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
75 FATEXP46	Em quantos dias a fadiga foi pior pela manhã?	Nenhum	1 dia	2 – 3 dias	4 – 5 dias	6 – 7 dias
76 FATEXP48	Com que frequência você percebeu que se cansava facilmente?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
77 FATEXP49	Com que frequência você pensou na fadiga?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
78 FATEXP50	Quanta fadiga você sentiu no dia em que ficou menos fatigado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
79 FATEXP51	Em média, com que facilidade você percebeu que estava ficando cansado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
80 FATEXP52	Até que ponto você ficou arrasado(a)/ prostrado(a)?	Nem um pouco	Um pouco	Mais ou menos	Muito	Muitíssimo
81 FATEXP54	Com que frequência você teve energia física?	Nunca	Raramente	Algumas vezes	Muitas vezes	Sempre
82 FATEXP56	Qual foi o nível da sua fadiga na maioria dos dias?	Nenhum	Leve	Moderado	Grave	Muito grave

