

# **Inconsistências lógicas em avaliações de estados de saúde definidos a partir do sistema EQ-5D: uma análise das preferências em saúde para a população de Minas Gerais**

**Carla de Barros Reis (CEDEPLAR/UFMG)**  
**Mônica Viegas Andrade (CEDEPLAR/UFMG)**  
**Kenya Noronha (CEDEPLAR/UFMG)**  
**Paul Kind (UNIVERSITY OF LEEDS)**

## **Resumo**

Inconsistências lógicas em preferências por estados de saúde ocorrem quando um estado logicamente pior, em termos de qualidade de vida, é preferível a um estado de saúde logicamente melhor. Este trabalho explora a consistência das valorações dos estados de saúde definidos pelo sistema descritivo EQ-5D realizadas pela população mineira. Comparam-se os níveis de inconsistências gerados em diferentes métodos de valoração; examina-se a associação entre as características individuais e inconsistências; e o efeito de se remover respondentes inconsistentes da estimação das tarifas sociais. Os resultados indicam a possibilidade de maior ocorrência de inconsistências em populações com maior heterogeneidade cultural e socioeconômica.

**Palavras-chave:** Inconsistência lógica; EQ-5D; Métodos de valoração; Estados de saúde.

## **Abstract**

Logical inconsistency for health states preferences occurs when one logically worse health state, in terms of quality of life, is ranked higher than a logically better health state. This study explores the consistency of the responses in a study of the EQ-5D health states valuations from a Brazilian population survey. It compares the level of inconsistency generated in different valuation methods; it examines the association between individual characteristics and the inconsistencies, and the effect of removing inconsistent respondents on the estimation of social value sets. The results indicate the possibility of higher occurrence of inconsistencies in populations with greater cultural and socioeconomic heterogeneity.

**Keywords:** Logical inconsistency; EQ-5D; Valuation methods; Health states.

**Área Temática 2: Economia**

# 1 INTRODUÇÃO

As avaliações que ponderam os custos monetários e os benefícios associados à introdução ou substituição de uma tecnologia em saúde, considerando as preferências sociais, são conhecidas como análises de custo-utilidade. Uma abordagem usual para se medir os benefícios neste tipo de análise é o uso dos Anos de Vida Ajustados pela Qualidade (AVAQ ou QALY do inglês *Quality-Adjusted Life Years*), que combina informações sobre mortalidade e qualidade de vida relacionada à saúde (PHILLIPS, 2009). A dimensão de qualidade de vida associada a distintas condições de saúde pode ser medida através de diversas técnicas, com o objetivo de gerar valores que reflitam a magnitude das preferências individuais. As principais técnicas utilizadas para a obtenção dos pesos são o Exercício da Troca de Tempo (do inglês *Time Trade-Off* - TTO), a Loteria Padrão (do inglês *Standard Gamble* - SG) e a Escala Analógica Visual – EAV (do inglês *Visual Analogue Scale* - VAS).

A percepção individual do estado de saúde geralmente é avaliada em diferentes domínios ou seja, em determinadas áreas do comportamento ou da condição humana, tais como as dimensões física, psicológica, econômica, espiritual e social. Um instrumento de avaliação de qualidade de vida que abarca alguns desses domínios é o sistema descritivo EuroQol 5 Dimensões (EQ-5D), desenvolvido por um grupo de especialistas da Universidade de York (*Measurement and Valuation of Health Group* – MVH) (GUDEX, 1994). O instrumento consiste em uma métrica que define qualidade de vida relacionada à saúde em termos de cinco dimensões (mobilidade, autocuidado, atividades usuais, dor/desconforto, ansiedade/depressão) com três níveis de severidade cada uma (sem problemas, alguns problemas e problemas extremos), resultando na descrição de 243 ( $3^5$ ) distintos estados de saúde. Dessa forma, cada estado de saúde é representado por uma seqüência de cinco algarismos, que correspondem aos níveis de severidade das respectivas dimensões. O estado de saúde 12113, por exemplo, reflete ausência de problemas nas dimensões de mobilidade, atividades usuais e dor/desconforto, alguns problemas na dimensão de autocuidado e problemas extremos de ansiedade/depressão.

Essa seqüência de algarismos apresenta uma escala ordinal interna definida pelo nível de severidade, de forma que, quanto maior a gravidade, pior é a condição de saúde da referida dimensão (DOLAN e KIND, 1996). Assim, possuir problemas extremos de mobilidade é considerado pior que possuir apenas alguns problemas de mobilidade que, por sua vez, é pior que não possuir qualquer problema para caminhar. Como os estados de saúde são formados pela combinação de elementos para cada uma das cinco dimensões, é possível para alguma destas descrições estabelecer uma relação ordinal esperada entre elas. Por exemplo, o estado de saúde 12231 é logicamente dominado pelo estado 12121, já que para cada dimensão do primeiro, o nível de severidade é igual ou maior (pior) que o correspondente grau de severidade do segundo. Valorações que violam este princípio são denominadas como logicamente inconsistentes (CRAIG e RAMACHANDRAN, 2006; DOLAN e KIND, 1996; BADIA, MONSERRAT e HERDMAN, 1999; DELVIN *et al.*, 2003; KRABBE, ESSINK-BOT e BONSEL, 1997). A inconsistência ocorre quando um estado de saúde logicamente dominado, em termos de qualidade de vida, recebe uma valoração superior àquela atribuída a um estado de saúde dominante.

Em estudos de estimação de parâmetros de preferências sociais por estados de saúde, a presença de respostas inconsistentes pode causar ruído nas estimativas. O tratamento dessas respostas ainda é uma questão em aberto na literatura (DEVLIN *et al.*, 2003), entretanto a censura de respostas inconsistentes tem sido empregada quando se opta pelo tratamento. No estudo de obtenção de pesos para a população canadense, por exemplo, dos 2394 respondentes que aceitaram participar da pesquisa, apenas 1145 (48%) são incluídos na análise primária. Foram excluídos os indivíduos que não completaram todas as avaliações, que atribuíram mesmo valor a todos os estados de saúde e que obtiveram mais que um par de estados de saúde com avaliações reversas à ordem lógica esperada (BANSBACK *et al.*, 2012). No estudo de estimação dos parâmetros dos estados de saúde definidos a partir do EQ-5D, para a população de Taiwan, Lee *et al.* (2013) excluem 38,79% dos respondentes da

amostra. Estes participantes cometeram pelo menos uma das seguintes inconsistências: todos os estados de saúde classificados como piores que a morte, atribuição de mesmo valor para todos os estados de saúde e, apresentar quatro ou mais inconsistências lógicas.

Dolan e Kind (1996) distinguem as inconsistências em primária e secundária. Enquanto a primeira decorre de limitações intrínsecas aos respondentes, a segunda refere-se ao método de valoração utilizado. Neste sentido, a comparação de diferentes métricas de valoração dos estados de saúde deve levar em consideração a capacidade de mensuração e de se produzir respostas logicamente consistentes. No caso da inconsistência primária é importante avaliar a frequência com que estas ocorrem em subgrupos populacionais específicos. A remoção dos respondentes logicamente inconsistentes pode alterar de forma significativa a representatividade amostral e, em último caso, os resultados de avaliações de tecnologias em saúde. Se por um lado, a remoção de dados inconsistentes pode comprometer a representatividade das preferências de grupos específicos da população, por outro lado, a inclusão dos respondentes que, a priori, oferecem informações inconsistentes, pode não captar as reais preferências em saúde destes grupos populacionais, além de piorar a performance da estimação das tarifas por estados de saúde.

Verificando os dados de Barcelona a partir de três métodos de obtenção de pesos, Badia, Monserrat e Herdman (1999) encontram percentuais de indivíduos inconsistentes iguais a 59.2%, 25.9% e 24.4%, para os métodos Time Trade-off, EAV e Ordenação, respectivamente. Quando se considera a razão entre os pares de estados de saúde inconsistentes e os pares potencialmente inconsistentes, o TTO também apresenta maior taxa média e mediana de violação da ordem lógica do EQ-5D. Enquanto a taxa média de inconsistências registra valor de quase 4% no método Time Trade-off, as taxas verificadas no Exercício de Ordenação e na Escala Analógica Visual são de 1.32% e 1.14%, respectivamente. Além disso, aumento da idade e baixos níveis de escolaridade são positivamente associados com altos níveis de inconsistência nos métodos cardinais, apesar de não haver qualquer efeito quando se considera o Exercício de Ordenação. Krabbe, Essink-Bot e Bonsel (1997) encontram uma taxa de inconsistência média de 4,6%, 4,3% e 2% para os métodos SG, TTO e EAV, respectivamente.

Utilizando dados do Reino Unido e Estados Unidos, Craig e Ramachandran (2006) encontram que respondentes inconsistentes são, em grande parte, idosos, mulheres, de níveis educacionais mais baixos, que reportaram ter tido dificuldades em executar as tarefas do TTO e da EAV e apresentam pior saúde auto-avaliada segundo o sistema descritivo EQ-5D. A remoção dos respondentes inconsistentes modificou os valores médios das preferências sociais obtidas através do método do TTO e da EAV para diversos estados de saúde. Estados de saúde brandos tiveram seus valores aumentados, enquanto estados severos apresentaram queda dos seus valores, aumentando a dispersão dos pesos médios. Bravata *et al.* (2005) encontram uma relação direta entre o tempo da entrevista e o nível de inconsistências. No estudo para Malásia, que investigou 152 pacientes de nove hospitais públicos, o método TTO foi considerado pelos respondentes mais difícil que a EAV, além de apresentar maior número de inconsistências lógicas: cerca de 63% dos respondentes foram totalmente consistentes na Escala Analógica Visual e apenas 15.7% deles foram no método Time Trade-off (YUSOF, GOH e AZMI, 2012).

Utilizando informações da EAV, no estudo para Nova Zelândia, Devlin *et al.* (2003) encontram um percentual de 21% de respondentes sem qualquer inconsistência em suas avaliações. O grupo de indivíduos que apresentaram, no máximo, uma inconsistência, formou a subamostra de maior proximidade com a amostra completa. Apesar da alta similaridade entre os coeficientes estimados para cada nível e dimensão de saúde, com mais de 2/3 das diferenças de médias absolutas inferiores a 0,05, e a igualdade da importância relativa nas duas amostras, os autores interpretam que os pesos estimados apresentaram diferenças significativas. Lamers *et al.* (2006), utilizando as avaliações dos estados de saúde definidos pelo EQ-5D de 309 indivíduos adultos da Holanda, encontram proporções iguais a 65% e 89% de respondentes inconsistentes na EAV e no TTO, respectivamente. A remoção dos

respondentes que apresentaram mais que três inconsistências, não resultaram em diferenças significativas dos coeficientes e pesos estimados na amostra completa.

Esse artigo tem como objetivo analisar em que medida as valorações dos respondentes respeitam a ordenação lógica esperada segundo três diferentes métodos de avaliação: Exercício de Ordenação, Escala Analógica Visual e método Time Trade-off. Além disso, examina-se a relação entre características individuais e a geração de inconsistências, além da sensibilidade dos parâmetros de valoração estimados em relação à remoção de respondentes inconsistentes. Utilizou-se uma pesquisa realizada no Brasil referente às estimações dos parâmetros de valorização dos estados de saúde, conduzida no estado de Minas Gerais, a partir do sistema descritivo EQ-5D (ANDRADE *et al.*, 2013). Este estudo a partir das preferências brasileiras abre espaço para a verificação de evidências quanto à qualidade das informações obtidas em avaliação de estados de saúde para países que apresentam maior heterogeneidade cultural e socioeconômica.

Os resultados mostram que apenas 8,7% dos pares de estados de saúde são avaliados de forma inconsistente no método TTO. Entretanto, as taxas de inconsistência associadas aos três métodos utilizados foram superiores aquelas encontradas na literatura internacional. O método Time Trade-off produziu maiores taxas de inconsistência que os demais métodos e a probabilidade de ser inconsistente é maior entre indivíduos com menor nível educacional e condição socioeconômica. Contudo, as preferências sociais são pouco afetadas após a exclusão dos indivíduos inconsistentes. Além desta introdução, este trabalho é composto por mais três seções. No capítulo 2 são apresentadas a base de dados e a definição de inconsistência lógica. No capítulo 3 seguem os resultados e, por fim, a discussão e as conclusões.

## **2 METODOLOGIA**

### **2.1 Base de Dados**

A base de dados utilizada neste trabalho foi obtida a partir dos resultados de uma pesquisa realizada pelo Centro de Desenvolvimento e Planejamento Regional - CEDEPLAR/UFMG, cujo objetivo era a estimação dos parâmetros de valorização dos estados de saúde em Minas Gerais a partir do sistema descritivo EQ-5D. A coleta de dados foi realizada entre outubro e dezembro de 2011 no estado de Minas Gerais. Foram investigados indivíduos alfabetizados, com idade entre 18 e 64 anos, residentes de áreas urbanas do estado. As entrevistas foram conduzidas de forma a selecionar apenas um indivíduo em cada domicílio e incentivos econômicos não foram fornecidos aos respondentes. O tamanho amostral foi definido de acordo com Censo Demográfico de 2010, de modo a obter margem de erro de 3%. No total, 3362 indivíduos foram investigados, dos quais 1115 eram residentes em Belo Horizonte, 626 na Região Metropolitana de Belo Horizonte e 1621 no interior do Estado. Considerou-se a técnica de amostragem probabilística, estratificada por sexo e idade, para todo o estado e para os três distintos níveis regionais.

As características demográficas e socioeconômicas da amostra são comparáveis às observadas em outras pesquisas domiciliares para o estado de Minas Gerais: Pesquisa por Amostra de Domicílios - PAD 2009 e a Pesquisa Nacional por Amostra de Domicílios - PNAD 2008 (TABELA 1). A versão brasileira do sistema descritivo EQ-5D foi culturalmente adaptada e o protocolo de entrevistas seguiu uma versão revisada (KIND, 2009) do estudo original do MVH (GUDEX, 1994). Dos 243 estados de saúde definidos pelo EQ-5D 102 estados foram diretamente valorados. Os estados foram agrupados em 26 blocos, cada um composto por seis perfis de saúde (dois brandos, dois moderados e dois severos). De acordo com Kind (2009), a classificação dos estados de saúde é realizada com base na métrica da distância do bloco, contabilizando-se a distância de cada uma das dimensões do estado de saúde analisado em relação ao índice de saúde perfeita (11111). Assim, as distâncias para os estados 22222 e

33333 são 5 e 10, respectivamente. A classificação segue a seguinte definição: i) *Brandos*: estados cujas distâncias estejam compreendidas entre 1 e 4; não possuam problemas de nível 3; e permitam, no máximo, 3 problemas de nível 2; ii) *Severos*: estados que não possuam problemas de nível 1 e que permitam, no mínimo, 2 problemas de nível 3, com exceção do pior estado de saúde hipotético (33333). iii) *Moderados*: demais estados de saúde, com exceção do estado de saúde perfeita (11111) e do pior estado de saúde hipotético (33333).

Devido ao tamanho da amostra, mais de 100 observações por estado de saúde foram garantidas, de acordo com as recomendações de Chuang e Kind (2010). Cada indivíduo um bloco contendo seis estados de saúde e o melhor e o pior estado de saúde (11111 e 33333, respectivamente), além da ‘Morte’, totalizando nove estados de saúde.

**TABELA 1 – Características sociodemográficas da amostra e comparação com bases de dados nacionais (%)**

Características	PPSUS	FJP	IBGE
Sexo			
Homem	51,58	52,43	52,08
Mulher	48,42	47,57	47,92
Faixa Etária			
18-34	43,3	47,23	46,2
35-49	33,95	32,5	33,37
50-59	16,25	15,29	15,9
>60	6,5	4,98	5,13
Nível Educacional			
Primário Incompleto	4,86	-	5,62
Primário Completo	24,37	-	7,7
Fundamental Completo	24,56	-	37,51
Médio Completo	37,64	-	36,43
Superior Completo	8,54	-	12,74
Estado Civil			
Casado	56,12	45,85	-
Viúvo	2,83	2,95	-
Divorciado ou Separado	7,92	6,71	-
Solteiro	33,13	44,41	-
Plano de Saúde			
Sim	31,36	28,34	35,38
Não	68,64	71,28	64,62
Saúde auto-avaliada			
Muito bom	25,35	29,17	31,18
Bom	52,01	49,41	48,99
Regular	20,49	18,32	17,12
Ruim	1,58	2,4	2,14
Muito ruim	0,49	0,65	0,57

Fonte: Andrade *et al.*, 2013.

O protocolo da entrevista inclui três instrumentos. No Instrumento A, ao indivíduo é solicitado que descreva seu próprio estado de saúde, em termos do sistema descritivo EQ-5D, e registre uma pontuação equivalente na Escala Analógica Visual, com as âncoras de 0 e 100 correspondendo ao pior e ao melhor estados de saúde. Depois, através do Exercício de Ordenação, ordenam-se os nove estados de saúde, segundo as preferências individuais. Ainda neste mesmo instrumento, realiza-se o preenchimento da EAV, indicando em que local da escala os participantes classificam os nove estados de saúde. Entre a aplicação do exercício de ordenação e o preenchimento da escala, os estados são apresentados de forma aleatória ao entrevistado, a fim de que a evidência de correlação entre os dois

métodos seja espontânea. É explicado que cada estado de saúde deve durar por 10 anos sem qualquer mudança, seguido por morte. Estes exercícios são realizados anteriormente ao método Time Trade-off com a finalidade de familiarizar os indivíduos às descrições dos perfis de saúde.

No Instrumento B, são apresentadas aos entrevistados escolhas entre alternativas que variam entre níveis de quantidade e qualidade de vida. Cada indivíduo valoriza sete estados de saúde (seis estados que abarcam os três níveis de severidade e o pior estado de saúde hipotético 33333), utilizando o método Time Trade-off através de um quadro de dupla-face, no qual em um dos lados se avaliam estados considerados pelos respondentes como melhores que a morte e, na outra face, estados registrados como piores que a morte. No primeiro caso, o respondente seleciona o número  $x$  de anos ( $x < 10$ ) sob o melhor estado de saúde hipotético (Saúde Perfeita) que equivale a viver por 10 anos no estado com algum decréscimo de saúde; quanto menor este período de equivalência, pior é o estado em questão. No caso dos estados piores que a morte, os indivíduos escolhem morrer imediatamente ou viver por um número  $x$  de anos ( $x < 10$ ) na condição específica de saúde, seguido por  $(10 - x)$  anos no estado de saúde perfeita como forma de compensação. Assim, quanto maior o tempo requerido no melhor estado de saúde para compensar um curto período de tempo vivido na condição em questão, pior é este estado (GUDEX, 1994). O Instrumento C investiga as características demográficas e socioeconômicas de cada respondente.

Como saúde integral e morte apresentam pontuações 1 e 0 respectivamente, então para estados classificados como melhores que a morte no TTO, seus valores são dados pela fórmula  $V = x/10$ , onde  $x$  é o período de tempo vivido com saúde integral. Para estados classificados como pior que a morte, seus valores são dados pela fórmula  $V = -x/(10-x)$ . Contudo, este procedimento gera valores com limite inferior igual a -19, enviesando para baixo os pesos. A exemplo de DOLAN *et al.* (1996), a fim de tratar esta distribuição assimétrica dos valores negativos, realizou-se uma transformação monotônica através da fórmula  $V_i = V/(1-V)$ , se  $V < 0$ . Dessa forma, estados classificados como piores que a morte foram limitados inferiormente pelo valor -1, trabalhando-se com um intervalo equidistante da morte em ambas direções, positiva e negativa. Esta transformação é comumente utilizada na literatura.<sup>1</sup> As pontuações obtidas a partir da Escala Analógica Visual também sofrem transformações, a fim de que os estados de saúde perfeita e morte representassem âncoras para todos os respondentes. Assim, segundo a fórmula  $EAV_j^t = \frac{EAV_j - EAV_{MORTE}}{EAV_{PERFEITA} - EAV_{MORTE}}$  obtém-se valorações padronizadas, de modo a fixar os valores das âncoras em 0 (Morte) e 1 (Saúde Perfeita).

## 2.2 Definição de inconsistência

O sistema descritivo EQ-5D consiste de uma sequência de cinco algarismos que representa as cinco dimensões de saúde, com três níveis de severidade cada uma. Essa sequência apresenta uma escala ordinal interna definida pelo próprio nível de severidade, de forma que, quanto maior a gravidade, pior é a condição de saúde da referida dimensão (DOLAN e KIND, 1996). Como os estados de saúde são formados pela combinação de elementos para cada uma das cinco dimensões, estas descrições podem ser classificadas de acordo com a relação ordinal esperada entre elas. Contudo, não são todos os pares de estados de saúde passíveis de comparação.

Um par lógico ou comparável é aquele que permite o estabelecimento de uma conexão lógica de dominância entre os estados de saúde. A dominância entre os estados de saúde se verifica quando a mudança em cada dimensão entre os pares de estados respeitar sempre o mesmo sentido. Por exemplo, o estado 12231 é logicamente dominado pelo estado 11221, já que para cada dimensão do primeiro, o nível de severidade é igual ou maior (pior) que o correspondente grau de severidade do segundo. Nesse caso ocorre mudança somente nas dimensões de cuidados pessoais e dor/desconforto. Transita-se da

---

<sup>1</sup> Para diferentes transformações dos pesos ver Lamers (2007).

presença de alguns problemas para lavar-se ou vestir-se para nenhum problema nos cuidados pessoais, e da presença de dor/desconforto extremos para a presença moderada de dores ou mal-estar. Por outro lado, a relação entre os estados 12231 e um outro estado, 21231, não pode ser logicamente determinada, dado que os níveis das dimensões de 12231 não são superiores ou inferiores aos correspondentes níveis de 21231, de forma monotônica. Enquanto na dimensão de mobilidade, o primeiro estado de saúde é preferível ao segundo, no domínio referente aos cuidados pessoais, o sentido da relação de preferência é contrário. Assim, como os respondentes podem atribuir pesos distintos as diferentes dimensões, torna-se difícil estabelecer uma relação de dominância entre as duas condições de saúde.

A inconsistência ocorre quando um estado de saúde logicamente dominado, segundo a definição pré-estabelecida anteriormente, recebe uma valoração superior àquela atribuída a um estado de saúde dominante, em termos de qualidade de vida. Desse modo, se para o indivíduo  $i$  o estado 31212 é preferível ao estado 11112, este par de estados é considerado inconsistente. Para que o indivíduo respeite a hipótese pré-definida de ordinalidade, o valor que atribui ao estado  $j$  deve ser maior que o valor atribuído ao estado  $k$ , quando  $k$  é considerado logicamente pior, no mínimo, em uma das dimensões e não melhor, nos demais domínios. Assim, considerando **monotonicidade estrita** na escolha do agente, a inconsistência na avaliação entre os pares de saúde  $j$  e  $k$  ocorre ou quando o valor que o indivíduo  $i$  atribui ao estado  $j$  é maior que àquele atribuído ao estado  $k$ , e simultaneamente, o estado de saúde  $j$  é dominado pelo estado  $k$ ; ou quando o valor atribuído ao estado  $j$  é menor que àquele atribuído ao estado  $k$ , e simultaneamente, o estado de saúde  $j$  domina o estado  $k$ . Definição alternativa que considera **monotonicidade não-estrita** na escolha do agente, também considera inconsistente o indivíduo que atribui um mesmo valor para estados de saúde comparáveis distintos.

### 2.3 Cálculo das inconsistências

A inconsistência interna será analisada utilizando-se dois indicadores: i) a porcentagem de respondentes que violam a ordem pré-estabelecida, no mínimo, em um par de estados de saúde; ii) a taxa de inconsistência individual, definida pela razão entre o número de pares inconsistentes e o número de pares comparáveis, ou seja, que possibilitam uma relação de dominância lógica, para cada indivíduo. Para ilustrar o cálculo da taxa de inconsistência, imagine que, utilizando um dado método de obtenção de preferências, um indivíduo produz o seguinte ordenamento de estados de saúde: 31113, 33333, 32113 e 22221. Existem duas inconsistências nesta avaliação, dado que o pior estado de saúde hipotético 33333 nunca pode ser considerado melhor que os estados 32113 e 22221. Quatro pares de estados de saúde são comparáveis, pois os estados 31113, 32113 e 22221 são preferíveis à condição 33333 e o estado 31113 é preferível à condição 32113. Os pares 31113 e 22221, e 32113 e 22221 não podem ser classificados como comparáveis segundo a definição apresentada. Dessa forma, o respondente realizou duas de quatro inconsistências possíveis, fornecendo uma taxa de inconsistência igual a 1/2. Esta taxa varia entre 0 (ausência de pares inconsistentes) e 1 (todos os pares lógicos são inconsistentes).

Através do conjunto de 243 estados de saúde possíveis gerados pelo sistema descritivo EQ-5D, pode-se formar  $C_{243,2} = 29.403$  pares, dos quais 7.533 (25,6%) são comparáveis por construção, ou seja, possuem relação de dominância lógica entre seus domínios (CRAIG e RAMACHANDRAN, 2006). Algumas inovações sugeridas pelo Grupo EuroQol (KIND, 2009) foram adotadas na pesquisa. Uma delas constitui-se na obtenção de 102 estados de saúde, de forma direta, e a utilização de 26 blocos fixos, cada um composto por seis estados de saúde pré-definidos, de forma a balancear os três níveis de severidade possíveis. Assim, cada estado de saúde somente é avaliado com aqueles estados que pertencem ao grupo em que se encontra. Este desenho impõe uma quantidade ainda menor de pares comparáveis, já que o número de estados de saúde expostos simultaneamente a uma mesma avaliação, é reduzido. Além disso, o cartão Morte não foi considerado no cálculo das inconsistências. Primeiro,

porque a Morte é uma das âncoras utilizadas no exercício Time Trade-off e, portanto, não apresenta avaliações. Segundo, e mais importante, permitir que um estado de saúde classificado como pior que a morte seja considerado como inconsistência viola o pressuposto básico do método baseado na troca.

Diante da importância de se considerar a qualidade e a representatividade dos dados em análise, verifica-se a possibilidade de diferenças na produção de inconsistências pelos três diferentes métodos de avaliação. Essas diferenças serão verificadas a partir de estatísticas de tendência central como média e mediana da taxa de inconsistência, além da proporção de indivíduos inconsistentes. Além disso, busca-se analisar se características individuais como educação, condição socioeconômica domiciliar e idade do respondente possuem algum efeito sobre a geração de escolhas inconsistentes. As diferenças por subgrupos de respondentes serão avaliadas, por meio das diferenças de médias das taxas de inconsistência e do modelo logit para respostas binárias. Diferenças de médias são obtidas a partir dos testes t de Student e os métodos não-paramétricos de Mann-Whitney e Kruskal-Wallis. Os testes t de Student e Mann-Whitney são usados para testar se duas amostras independentes foram extraídas de populações com médias iguais. Contudo, o segundo serve como alternativa ao primeiro, principalmente, quando a variável populacional analisada não segue uma distribuição normal e/ou as amostras forem pequenas. Quanto ao teste Kruskal-Wallis, este é extremamente útil para decidir se k amostras ( $k > 2$ ) independentes provêm de populações com médias iguais (SIEGEL e CASTELLAN Jr., 2006).

Na regressão logística, a variável resposta dicotômica  $y_i$  assume valor igual à unidade se o indivíduo  $i$  avalia, no mínimo, um par de estados de saúde de forma a violar a ordenação lógica pré-estabelecida; e zero, caso o respondente seja totalmente consistente em suas decisões. Portanto, as probabilidades de sucesso estão associadas a agentes inconsistentes. As variáveis independentes, listadas na TABELA 2, referem-se às características individuais dos respondentes, além de variáveis relativas ao procedimento de pesquisa. A variável socioeconômica utilizada foi o “Critério Brasil”, um instrumento de segmentação econômica que diferencia a população de acordo com dois fatores: o grau de instrução do chefe de família e as características domiciliares, que são definidas pela posse de itens, tais como, geladeira, televisão em cores, automóvel, empregada mensalista, etc. O critério atribui pontos em função de cada uma destas características e a soma destes pontos – que varia entre 0 e 46 – define em qual “classe econômica” o domicílio se enquadra (ABEP, 2013). Há uma correspondência entre as faixas de pontuação do critério e estas classes, que são definidas por A1, A2, B1, B2, C1, C2, D e E, sendo a classe A1 aquela que abrange os domicílios com pontuação entre 42 e 46 (mais ricos) e a classe E a que abarca os domicílios entre 0 e 7 pontos (mais pobres). Quatro grupos sínteses são construídos a partir das classes socioeconômicas, a saber, os grupos A, B, C, D-E.

Por fim, excluem-se os indivíduos inconsistentes da amostra e calculam-se as tarifas médias para todos os estados de saúde. Considerando a possibilidade de um erro aleatório na avaliação individual, acredita-se ser razoável que um indivíduo cometa apenas uma inconsistência. Neste sentido, também se considera consistente o grupo de respondentes com no máximo uma violação da ordem lógica. No caso do TTO e da EAV, obtém-se os valores observados médios. A fim de se observar o efeito da exclusão de indivíduos inconsistentes da amostra sobre os coeficientes utilizados para a interpolação de estados de saúde que não foram observados diretamente, são calculados os pesos médios para cada estado de saúde, considerando o modelo de efeitos aleatórios. A variável dependente corresponde ao valor da desutilidade ( $I - TTO/I - VAS$ ) associada a um determinado estado de saúde e o vetor de variáveis independentes - TABELA 3, é composto por 10 variáveis explicativas binárias correspondentes às dimensões do EQ-5D, que tomam valor igual a 1, se apresentam determinando nível de severidade e zero, caso contrário<sup>2</sup>.

---

<sup>2</sup> Para detalhes do processo de estimação dos pesos TTO através do modelo de efeitos aleatórios, ver Andrade *et al.*, 2013.



**TABELA 2 – Descrição e Especificação das Variáveis do Modelo de Regressão Logística**

Variáveis	Descrição	Especificação
Inconsistente	Variável dependente relacionada à inconsistência do respondente em suas avaliações	1 se o indivíduo gera pelo menos uma inconsistência 0 se o indivíduo é consistente
Educação	Nível de escolaridade do respondente: Primário Incompleto Primário Completo Fundamental Completo Médio Completo Superior Completo	1 se Primário Incompleto (referência) 2 se Primário Completo 3 se Fundamental Completo 4 se Médio Completo 5 se Superior Completo
Idade	Idade do respondente em anos	contínua
Classe Socioeconômica	Classe socioeconômica do domicílio do respondente: Classe A Classe B Classe C Classe D-E	Pontuação no Critério Brasil: 35 a 46 pontos - Classes A1 e A2 (referência) 23 a 34 pontos - Classes B1 e B2 14 a 22 pontos - Classes C1 e C2 0 a 13 pontos - Classes D e E
Homem	Sexo do respondente	1 se homem; 0 se mulher (referência)
Estado Civil	Estado civil do respondente Casado Viúvo Divorciado/separado Solteiro	1 se casado (referência) 2 se viúvo 3 se divorciado/separado 4 se solteiro
Trabalhou na área da saúde	Experiência de trabalho do respondente na área da saúde	1 se trabalhou; 0 se não trabalhou (referência)
Possui plano de saúde	Posse de plano de saúde pelo respondente	1 se possui; 0 se não possui (referência)
Saúde auto-avaliada	Como o respondente avalia sua própria saúde: Muito bom ou bom Regular Ruim ou muito ruim	1 se muito bom ou bom (referência) 2 se regular 3 se ruim ou muito ruim
Feliz	Como o respondente avalia seu nível de felicidade: Feliz ou muito feliz Infeliz ou não muito feliz	1 se feliz ou muito feliz 0 se infeliz ou não muito feliz (referência)
Tempo de Entrevista	Tempo de duração da entrevista	contínua
Dificuldade no TTO	Como o respondente avalia o nível de dificuldade do método da Troca de Tempo	1 se difícil ou muito difícil 0 se fácil ou muito fácil (referência)

Fonte: elaboração dos autores.

Quanto ao Exercício de Ordenação, a estimação de pesos cardinais a partir de informações ordinais se dá a partir do modelo logit condicional, comumente utilizado na literatura (SALOMON, 2003; CRAIG, BUSSCHBACH e SALOMON, 2009; McCABE *et al.*, 2006). A variável resposta corresponde à posição do estado de saúde, de modo que a primeira posição refere-se a melhor escolha, segundo as preferências individuais. As variáveis independentes também estão associadas às dimensões e aos níveis de severidade do sistema EQ-5D, como mostra a TABELA 3. Função de avaliação análoga ao modelo previamente estimado (ANDRADE *et al.*, 2013) a partir dos pesos obtidos pelo método do Time Trade-Off foi adotada para facilitar a comparação. Sabe-se que os AVAQS são ancorados pelo estado de saúde perfeita (1) e pela morte (0), de modo a permitir comparações de intervenções de saúde distintas em uma escala comum. Contudo, o modelo logit condicional produz valores estimados definidos arbitrariamente em uma escala, de acordo com as hipóteses de identificação do modelo. Neste sentido, são realizadas duas diferentes normalizações dos valores a fim de que os pesos estimados para saúde perfeita e morte igualem a 1 e 0, respectivamente. Enquanto na normalização Rank B estabeleceu-se um pareamento entre o valor do pior estado de saúde (33333) e sua utilidade estimada através do método TTO, no Rank C atribui-se utilidade igual a zero para o estado morte<sup>3</sup>.

Realizam-se comparações entre os pesos da amostra completa e aqueles resultantes da amostra que exclui os indivíduos com mais de uma inconsistência em suas avaliações. O cálculo do número de

<sup>3</sup> Para mais informações com respeito aos procedimentos de normalização, ver Salomon (2003).

inconsistências e a construção das taxas de inconsistência por indivíduo foram realizados através do software livre C++. Todas as análises estatísticas são computadas usando o programa Stata 12.0.

**TABELA 3 – Descrição e Especificação das Variáveis do Modelo Logit Condicional (Exercício de Ordenação) e do Modelo de Efeitos Aleatórios (Time Trade-Off e Escala Analógica Visual)**

Variáveis	Descrição	Especificação
MO2	Estado de saúde apresenta problema moderado na dimensão mobilidade	1 se mobilidade é nível 2; 0 caso contrário
MO3	Estado de saúde apresenta problema grave na dimensão mobilidade	1 se mobilidade é nível 3; 0 caso contrário
SC2	Estado de saúde apresenta problema moderado na dimensão auto-cuidado	1 se auto-cuidado é nível 2; 0 caso contrário
SC3	Estado de saúde apresenta problema grave na dimensão auto-cuidado	1 se auto-cuidado é nível 3; 0 caso contrário
UA2	Estado de saúde apresenta problema moderado na dimensão atividades habituais	1 se atividades habituais é nível 2; 0 caso contrário
UA3	Estado de saúde apresenta problema grave na dimensão atividades habituais	1 se atividades habituais é nível 3; 0 caso contrário
PD2	Estado de saúde apresenta problema moderado na dimensão dor/desconforto	1 se dor/desconforto é nível 2; 0 caso contrário
PD3	Estado de saúde apresenta problema grave na dimensão dor/desconforto	1 se dor/desconforto é nível 3; 0 caso contrário
AD2	Estado de saúde apresenta problema moderado na dimensão ansiedade/depressão	1 se ansiedade/depressão é nível 2; 0 caso contrário
AD3	Estado de saúde apresenta problema grave na dimensão ansiedade/depressão	1 se ansiedade/depressão é nível 3; 0 caso contrário

Fonte: elaboração dos autores.

### 3 RESULTADOS

Os resultados encontrados para o caso brasileiro representam um baixo índice de inconsistências, já que apenas 8,7% dos pares de estados comparáveis no exercício TTO são inconsistentes, quando se considera o conceito de monotonicidade estrita das escolhas. A TABELA 4 apresenta as estatísticas sobre inconsistências lógicas para todos os métodos. Apesar de se observar uma distribuição aproximadamente equânime na proporção de indivíduos que cometem alguma inconsistência entre os métodos, a taxa média e mediana de inconsistência foi superior no Time Trade-off. Verifica-se que, enquanto no TTO, 8,7% de todos os pares de estados de saúde que possuem ordenamento lógico possível são inconsistentes, cerca de 6,0% e 4,9% dos pares que poderiam ser considerados inconsistentes são, de fato, classificados desta forma, no Exercício de Ordenação e na Escala Analógica Visual, respectivamente. Estes resultados indicam a existência de inconsistências secundárias, ou seja, aquelas derivadas, em grande parte, por algum aspecto do procedimento de medida.

Quando se considera o conceito de monotonicidade não-estrita, tanto a porcentagem de indivíduos inconsistentes, quanto às taxas médias e medianas de inconsistência lógica são menores para o Exercício de Ordenação se comparado aos demais métodos. Cerca de 86% dos respondentes produziram, no mínimo, uma inconsistência ao realizarem a tarefa do TTO, enquanto cerca de 63% e

67% dos indivíduos violaram a ordem lógica pré-estabelecida na ordenação e na EAV, respectivamente. Além disso, no método Time Trade-off, existe um maior número de respondentes com taxa de inconsistência superior a 50% - 228 comparado a 7 e a 17 na ordenação e na EAV, respectivamente. Taxas de inconsistências médias são maiores que as medianas, sugerindo que alguns indivíduos altamente inconsistentes estão enviesando a média para cima. Quanto à taxa de inconsistência, verifica-se que, enquanto no TTO, 21,5% de todos os pares de estados de saúde que possuem ordenamento lógico possível são inconsistentes, cerca de 7,5% e 8,5% dos pares que poderiam ser considerados inconsistentes são, de fato, classificados desta forma no Exercício de Ordenação e na Escala Analógica Visual, respectivamente.

**TABELA 4 – Inconsistências lógicas por métodos de avaliação**

	monotonicidade estrita			monotonicidade não-estrita		
	Ordenação	EAV	TTO	Ordenação	EAV	TTO
% respondentes inconsistentes	53,46	47,02	51,72	62,80	66,90	86,30
Taxa de Inconsistência Lógica						
Média	0,060	0,049	0,087	0,075	0,085	0,215
Mediana	0,045	0,000	0,067	0,050	0,053	0,182

Fonte: Elaboração dos autores.

No caso da monotonicidade estrita, no qual não se consideram inconsistentes os indivíduos que oferecem mesmo valor ou posição a estados de saúde comparáveis diferentes, a dominância de inconsistências entre métodos se altera, com a Escala Analógica Visual apresentando um menor nível de violações. Estas diferenças evidenciam a frequência de “coincidências” entre os valores obtidos nos métodos cardinais vis-à-vis a ordenação. O TTO e a EAV, por serem medidas cardinais, sugerem um julgamento mais refinado e, portanto, uma maior diferenciação entre estados de saúde com descrições similares. Contudo, verifica-se a situação oposta, em que os instrumentos cardinais apresentam maiores frequências de valores idênticos para situações de saúde distintas.

A TABELA 5 revela o percentual de respondentes por número de inconsistências apresentadas e sua distribuição acumulada. Percebe-se que os respondentes que violam a ordem lógica não representam as exceções. Quando não se considera inconsistência avaliar distintos estados de saúde com pesos iguais, o número máximo de inconsistências ocorre na EAV, que apresenta até 13 pares de estados comparáveis inconsistentes. No entanto, mais que 90% da amostra cometem até três inconsistências em todos os métodos. No caso da monotonicidade não-estrita, apenas cerca de 1/3 da amostra é totalmente consistente em suas escolhas nos Exercícios de Ordenação e Escala Analógica Visual, enquanto este percentual chega a menos que 15%, no caso do método Time Trade-off. Os números máximos de inconsistências crescem de forma expressiva na EAV e no TTO, chegando a 20 e a 15 pares, respectivamente.

A TABELA 6 apresenta a taxa média de inconsistência lógica para cada método, de acordo com as características socioeconômicas e demográficas dos respondentes. Um maior número de anos de educação formal parece estar negativamente associado ao aumento das inconsistências, sendo que esta relação apresenta-se de forma monotônica para o método Time Trade-off, considerando monotonicidade não-estrita. O aumento da idade parece estar positivamente correlacionado com a elevação das inconsistências no método da TTO, mas parece não possuir efeitos nas demais métricas. As variáveis referentes à condição socioeconômica e a posse de plano de saúde também estão associadas, de forma monotônica e significativa, ao número de inconsistências. Se o indivíduo não possui plano de saúde ou pertence às classes socioeconômicas mais empobrecidas, a taxa de inconsistência é maior.

**TABELA 5 – Percentual de respondentes inconsistentes (%) e distribuição acumulada, por método**

Pares Inconsistentes	monotonicidade estrita						monotonicidade não-estrita					
	Ordenação		EAV		TTO		Ordenação		EAV		TTO	
	(%)	F	(%)	F	(%)	F	(%)	F	(%)	F	(%)	F
0	46,54	46,54	52,98	52,98	48,31	48,31	37,23	37,23	33,01	33,01	13,72	13,72
1	23,32	69,86	22,96	75,94	23,21	71,52	25,23	62,47	26,04	59,05	21,44	35,16
2	13,89	83,75	11,52	87,46	13,57	85,09	15,77	78,24	16,04	75,10	19,77	54,93
3	7,03	90,78	5,48	92,94	7,67	92,75	9,22	87,46	10,21	85,30	16,56	71,49
4	4,04	94,82	3,11	96,05	3,05	95,81	5,39	92,85	5,81	91,11	11,17	82,66
5	2,16	96,98	1,74	97,79	1,86	97,66	3,02	95,87	3,98	95,09	5,78	88,44
6	1,38	98,35	1,02	98,80	1,32	98,98	1,98	97,84	1,92	97,01	4,22	92,66
7	0,81	99,16	0,60	99,40	0,51	99,49	1,11	98,95	1,02	98,02	2,73	95,39
8	0,24	99,40	0,18	99,58	0,30	99,79	0,36	99,31	0,69	98,71	1,35	96,74
9	0,30	99,70	0,09	99,67	0,15	99,94	0,30	99,61	0,42	99,13	0,90	97,63
10	0,15	99,85	0,15	99,82	0,06	100,00	0,24	99,85	0,30	99,43	0,57	98,20
11	0,06	99,91	0,12	99,94			0,06	99,91	0,21	99,64	0,45	98,65
12	0,09	100,00	0,03	99,97			0,09	100,00	0,09	99,73	0,30	98,95
13			0,03	100,00					0,15	99,88	0,60	99,55
14									.	.	0,09	99,64
15									0,03	99,91	0,36	100,00
16									0,03	99,94		
19									0,03	99,97		
20									0,03	100,00		

Fonte: Elaboração dos autores.

Pessoas viúvas, que se autoavaliam infelizes ou não muito felizes e que classificam a própria saúde como regular, ruim ou muito ruim apresentam maiores taxas médias de inconsistências. Sexo parece não estar associado ao nível de pares logicamente reversos na ordenação. Os resultados apresentados na TABELA 7 corroboram as evidências que, de fato, o nível educacional e a condição socioeconômica são determinantes importantes na geração de pares inconsistentes. Contudo, a importância destes elementos varia de acordo com a métrica utilizada na obtenção das preferências. Escolaridade possui efeitos negativos sobre inconsistências lógicas, tanto na ordenação quanto no método Time Trade-off, ao passo que na EAV, este efeito é positivo. Enquanto a diminuição das inconsistências só se faz sentir a partir do ensino médio e superior para o exercício de Ordenação, esta relação é extremamente forte no método TTO, de modo que qualquer ascensão no grau de escolaridade possui reflexos significativos na queda da inconsistência. Indivíduos com ensino superior completo possuem aproximadamente 60% a menos de chances de cometerem inconsistência, relativo aos respondentes que não completaram o ensino primário.

A idade parece possuir algum efeito somente no TTO. Contudo, a magnitude da diminuição das chances de produção de inconsistências é bastante irrisória. No que se refere ao nível socioeconômico, quanto pior as condições materiais do domicílio, maiores as chances de se cometer ordenações de forma a violar a ordem lógica, em todos os métodos. Os indivíduos cujos domicílios se encaixam nas classes D ou E possuem até 141% a mais de chances de cometer inconsistências que indivíduos residentes nos domicílios mais abastados. Quanto às demais variáveis, um resultado contra-intuitivo se faz presente: apresentar dificuldades na execução das tarefas do *TTO* diminui a probabilidade de ser inconsistente nos métodos cardinais. Talvez, este resultado possa ser explicado pelo alto nível de esforço e empenho que os indivíduos impuseram no desenvolvimento destas tarefas a fim de compreender a natureza do jogo, diferentemente dos respondentes desinteressados ou que não entenderam o objetivo dos instrumentos e que, portanto, resultaram em maiores números de

inconsistências. Tempo de duração da entrevista e possuir plano de saúde parecem não apresentar qualquer efeito.

**TABELA 6 – Taxa média de inconsistência lógica por características sociodemográficas e de saúde**

Características Individuais	monotonicidade estrita			monotonicidade não-estrita		
	Ordenação	EAV	TTO	Ordenação	EAV	TTO
<b>Educação</b>						
Primário incompleto	0.071	0.045	0.111	0.083	0.080	0.252
Primário completo	0.065	0.058	0.095	0.080	0.094	0.238
Fundamental completo	0.069	0.055	0.100	0.086	0.089	0.224
Médio completo	0.055	0.044	0.076	0.069	0.079	0.203
Superior completo	0.038	0.039	0.058	0.047	0.070	0.161
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
<b>Faixa etária</b>						
18-25	0.060	0.047	0.089	0.072	0.076	0.199
26-35	0.060	0.053	0.087	0.078	0.094	0.212
36-45	0.064	0.053	0.084	0.079	0.090	0.224
46-55	0.061	0.047	0.089	0.075	0.081	0.229
55-64	0.053	0.047	0.083	0.066	0.080	0.221
	$p < 0.001$	$p < 0.001$	$p = 0.012$	$p < 0.001$	$p < 0.001$	$p < 0.001$
<b>Classe socioeconômica</b>						
A	0.035	0.025	0.069	0.049	0.061	0.185
B	0.052	0.044	0.075	0.067	0.077	0.199
C	0.066	0.055	0.093	0.080	0.090	0.221
D-E	0.080	0.060	0.124	0.092	0.097	0.281
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
<b>Possui Plano de Saúde</b>						
Sim	0.051	0.044	0.078	0.066	0.079	0.195
Não	0.064	0.052	0.091	0.079	0.087	0.225
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
<b>Sexo</b>						
Homem	0.059	0.049	0.087	0.074	0.086	0.219
Mulher	0.061	0.051	0.086	0.075	0.083	0.209
	$p = 0.257$	$p = 0.059$	$p = 0.337$	$p = 0.086$	$p = 0.011$	$p < 0.001$
<b>Estado Civil</b>						
Solteiro	0.060	0.047	0.090	0.075	0.080	0.210
Casado	0.058	0.051	0.084	0.073	0.087	0.215
Divorciado	0.065	0.051	0.075	0.078	0.089	0.218
Viúvo	0.073	0.045	0.129	0.087	0.083	0.268
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
<b>Saúde auto-avaliada</b>						
Bom ou Muito bom	0.057	0.046	0.085	0.073	0.082	0.208
Regular	0.069	0.062	0.095	0.081	0.095	0.239
Ruim ou Muito ruim	0.094	0.065	0.098	0.103	0.107	0.233
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
<b>Felicidade</b>						
Feliz ou Muito feliz	0.058	0.048	0.086	0.072	0.083	0.214
Infeliz ou Não muito feliz	0.074	0.060	0.091	0.091	0.097	0.222
	$p < 0.001$	$p < 0.001$	$p = 0.017$	$p < 0.001$	$p < 0.001$	$p = 0.792$

Fonte: elaboração dos autores.

Nota: Os p-valores apresentados foram estimados utilizando os teste Kruskal-Wallis e Mann-Whitney. O teste t oferece resultados semelhantes.

Sexo possui efeitos de pequena magnitude, mas significativos. Enquanto homens apresentam 9,2% a mais de chances de cometerem inconsistências na ordenação, as mulheres possuem quase 5% a mais de chances de gerarem ordenações reversas no TTO. Ter experiência de trabalho na área da saúde e se perceber feliz ou muito feliz diminuem as chances de se cometer inconsistência, principalmente, na Ordenação e na EAV. Saúde auto-avaliada possui efeitos somente na Escala e os viúvos possuem cerca de 42% a mais de chances de cometer julgamentos inconsistentes no método TTO, em relação aos casados.

Quando se considera o conceito de monotonicidade não-estrita, os efeitos da saúde auto-avaliada sobre as inconsistências passam a ser significativos, com os indivíduos que dizem possuir saúde regular, ruim ou muito ruim apresentando uma redução nas chances de se cometer inversões na ordenação, em relação aos respondentes que percebem sua saúde como boa ou muito boa. No que se refere à métrica TTO, possuir plano de saúde reduz as chances de se cometer inconsistências, enquanto ser feliz apresenta efeito oposto. Quando as avaliações são realizadas a partir da EAV, o aumento da idade eleva, de forma significativa, esta chance.

**TABELA 7 – Razão de chances das inconsistências lógicas por características sociodemográficas e de saúde**

Características Individuais	monotonicidade estrita					
	Ordenação		EAV		TTO	
	Odds	dp	Odds	dp	Odds	dp
Educação						
Primário completo	1,019	0,066	1,521***	0,098	0,678***	0,045
Fundamental completo	1,054	0,069	1,253***	0,082	0,754***	0,051
Médio completo	0,891*	0,059	1,126*	0,075	0,539***	0,037
Superior completo	0,657***	0,053	1,212**	0,097	0,432***	0,035
Idade	0,999	0,001	0,998	0,001	0,994***	0,001
Classe socioeconômica						
B	1,509***	0,095	1,669***	0,109	0,833***	0,051
C	1,672***	0,109	2,071***	0,139	0,909	0,058
D-E	2,413***	0,202	2,162***	0,182	1,178**	0,097
Homem	1,092***	0,027	1,004	0,024	0,955*	0,023
Estado Civil						
Viúvo	1,034	0,083	0,858*	0,068	1,423***	0,116
Divorciado/Separado	0,979	0,046	0,896**	0,042	0,832***	0,039
Solteiro	0,970	0,028	0,948*	0,027	1,103***	0,031
Trabalhou na área da saúde	0,875***	0,032	0,871***	0,032	1,156***	0,043
Possui Plano de Saúde	0,999	0,027	0,997	0,027	0,985	0,026
Saúde auto-avaliada						
Regular	1,024	0,033	1,229***	0,039	1,064*	0,035
Ruim ou Muito ruim	0,983	0,091	0,844*	0,077	1,137	0,104
Feliz	0,775***	0,029	0,811***	0,031	0,900***	0,034
Tempo de Entrevista	1,000	0,000	1,001***	0,000	1,000	0,000
Dificuldade no TTO	1,104***	0,027	0,841***	0,020	0,836***	0,020
Constante	0,873	0,099	0,535***	0,061	2,730***	0,310

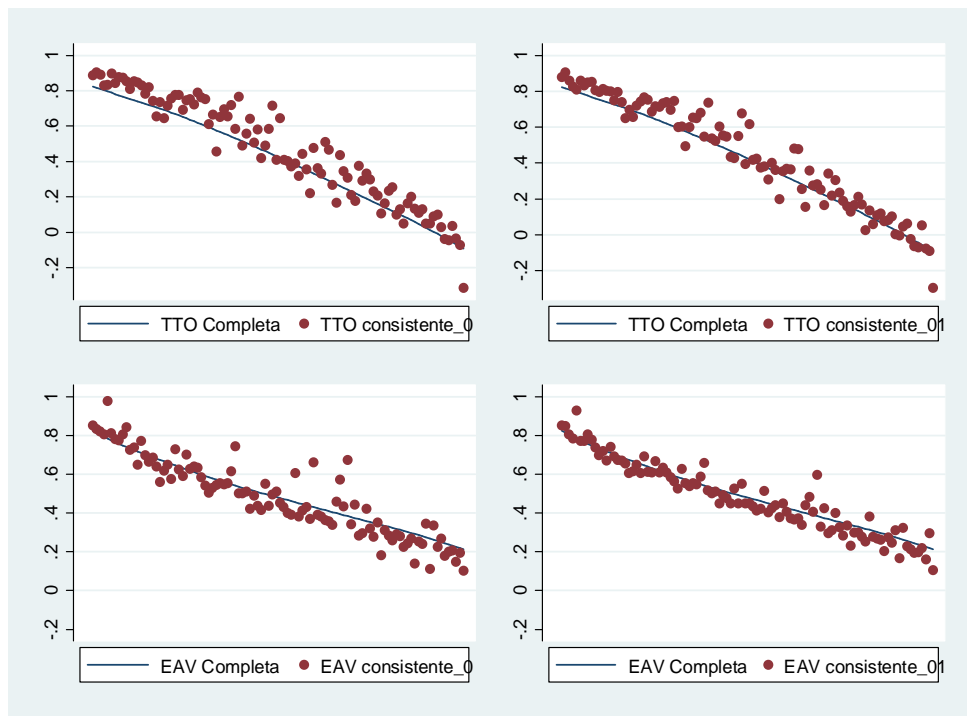
Fonte: Elaboração dos autores.

\*\*\* 1% de significância \*\* 5% de significância \* 10% de significância

Estes resultados mostram evidências da presença de inconsistências primárias, ou seja, àquelas geradas por limitações dos respondentes. Sendo assim, se os julgamentos reversos são mais frequentes nos subgrupos populacionais de menor escolaridade ou em pior situação socioeconômica, a exclusão de respostas inconsistentes do cálculo final das tarifas pode levar, no limite, à geração de médias que sub-representam os gostos destes grupos populacionais. Diferenças sutis são percebidas, no que se refere aos efeitos das inconsistências lógicas sobre as tarifas dos estados de saúde. As figuras 1 e 2 apresentam as comparações dos pesos médios obtidos de subamostras consistentes, ou seja, amostras de indivíduos com zero e com no máximo uma inconsistência, tendo como base a amostra com todos os indivíduos participantes. Considerando-se o conceito de monotonicidade estrita das escolhas na definição de inconsistência, as subamostras de respondentes sem qualquer inconsistência (TTO\_0) ou com, no máximo, uma inconsistência (TTO\_01), apresentam pesos médios observados do TTO superiores àqueles encontrados para a amostra completa, principalmente para os estados de saúde brandos e moderados. Já na EAV, os valores médios observados das referidas subamostras apresentam-se ligeiramente inferiores aos encontrados na amostra completa, principalmente para os estados severos.

Quando se consideram os pesos preditos para o Exercício de Ordenação, verifica-se grande pareamento entre as médias amostrais no Rank B. No que se refere à forma de normalização que atribui utilidade zero ao estado Morte - Rank C, as amostras consistentes apresentam valores médios suavemente inferiores aos correspondentes valores associados à amostra completa. Quando se considera como inconsistentes os indivíduos que, em um determinado par de estados de saúde, atribuem valores maiores ou iguais a estados dominados, os pesos médios se comportam de forma semelhante em todas as subamostras analisadas.

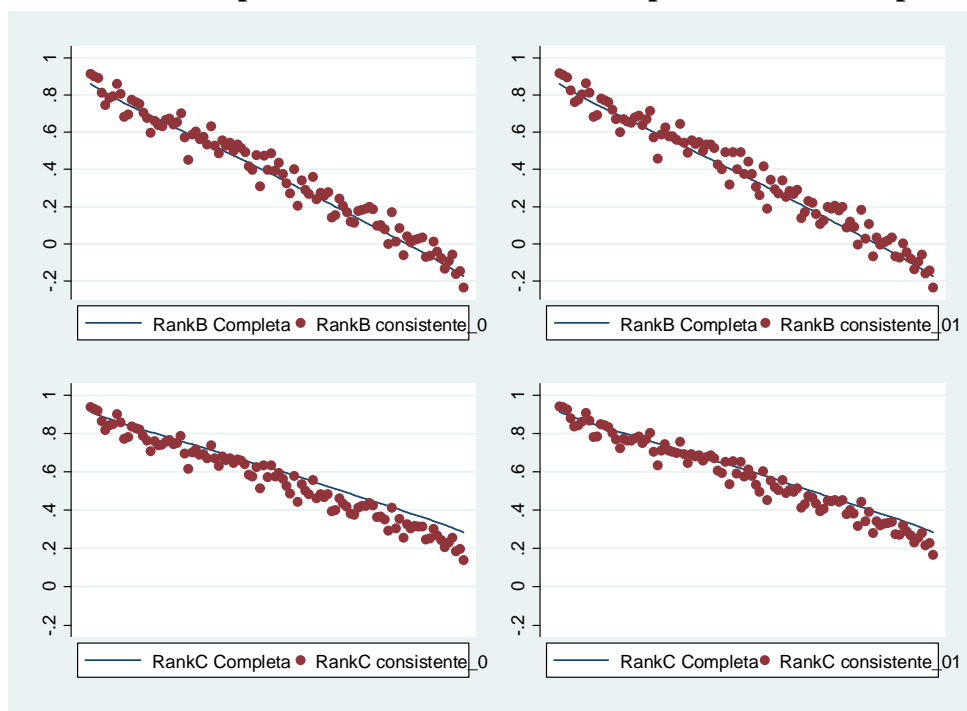
**FIGURA 1 – Valores médios observados do TTO e da EAV para amostra completa e consistente**



Fonte: Elaboração dos autores.

Nota: TTO<sub>i</sub> refere-se à amostra com respondentes que apresentaram um número *i* de inconsistências. Assim, TTO<sub>0</sub> e TTO<sub>01</sub> representam as amostras de indivíduos com 0 e 0-1 inconsistências no TTO, respectivamente.

**FIGURA 2 - Valores médios preditos do Rank B e Rank C para amostra completa e consistente**



Fonte: Elaboração dos autores.

Nota: RankC<sub>i</sub> refere-se à amostra com respondentes que apresentaram um número *i* de inconsistências. Assim, RankC<sub>0</sub> e RankC<sub>01</sub> representam as amostras de indivíduos com 0 e 0-1 inconsistências no Rank C.

De acordo com a TABELA 8, em todos os métodos, a subamostra de indivíduos com, no máximo, uma inconsistência, apresentou um conjunto de pesos bastante similar ao da amostra completa. O coeficiente de *Pearson* foi de 0,989 para os pesos obtidos pelo método EAV da amostra completa e da subamostra de respondentes com zero ou uma inconsistência. No que se refere aos dados derivados a partir da Ordenação, para a referida subamostra, as correlações são quase perfeitas.

**TABELA 8 – Correlação de Pearson entre os pesos da amostra completa e das subamostras consistentes**

Inconsistências	TTO	EAV	Rank B	Rank C
<b>0</b>	0,987	0,964	0,998	0,998
<b>0-1</b>	0,995	0,989	0,999	0,999

Fonte: Elaboração dos autores.

Nota: considera-se monotonicidade estrita na definição de inconsistências. Os coeficientes de *Pearson* para o caso de monotonicidade não-estrita das escolhas seguem padrões semelhantes, com menores níveis de correlação.

Se considerarmos a censura de dados inconsistentes, comumente realizada na literatura, os achados encontrados neste trabalho revelam que, para o caso de Minas Gerais, a exclusão destas informações afetaria muito sutilmente os pesos populacionais dos estados de saúde. Talvez, este achado seja resultante da grande proporção de indivíduos consistentes ou com apenas uma inconsistência na amostra completa, considerando-se monotonicidade estrita nas escolhas. A fim de se observar o efeito da exclusão de indivíduos inconsistentes da amostra sobre os coeficientes utilizados para a interpolação de estados de saúde que não foram observados diretamente, a TABELA 9 apresenta os resultados do modelo logit condicional, segundo a forma de normalização Rank B, que apresentou melhor ajuste, e os resultados do modelo de efeitos aleatórios para o método Time Trade-off e Escala Analógica Visual,



utilizando tanto a amostra completa quanto a amostra censurada, que exclui os indivíduos com mais de uma inconsistência em suas avaliações.

No caso do TTO, a maior diferença entre os coeficientes se deu na presença de alguns problemas na dimensão de mobilidade. Os coeficientes apresentaram pequenas diferenças ao se excluir indivíduos com maior número de inconsistências. Contudo, a importância relativa das dimensões se alterou: problemas moderados de mobilidade passam a gerar menor desutilidade que problemas com os cuidados pessoais, e problemas extremos com as atividades habituais geram menores prejuízos, em termos de qualidade de vida, que dor/desconforto extremos. As diferenças absolutas médias, as correlações entre os pesos e os números de estados de saúde cujas diferenças absolutas entre pesos foram superiores a 0.05 e 0.10 sofreram aumentos sutis em seus valores.

Quanto às tarifas estimadas com base nos dados ordenados e da EAV, apesar dos coeficientes apresentarem maiores diferenças ao se excluir indivíduos com maior número de inconsistências, a importância relativa das dimensões permaneceu inalterada. Os problemas severos na dimensão de mobilidade continuam gerando os maiores decréscimos de utilidade. Em ambos os grupos, depois das complicações na dimensão de mobilidade, problemas ligados aos cuidados pessoais possuem maior importância que dificuldades na realização das atividades habituais, seguidas por dor ou desconforto e, finalmente, ansiedade ou depressão. Além da permanência da alta correlação linear, a diferença absoluta média permanece inferior a 10%.

**TABELA 9 – Coeficientes estimados a partir do modelo logit condicional (Ordenação) e do modelo de efeitos aleatórios (Troca de Tempo e Escala Analógica Visual) para as amostras completa e consistente**

Variáveis	TTO	TTO_01	VAS	VAS_01	Rank B	Rank B_01
<i>Mobilidade 2</i>	0,128***	0,109***	0,150***	0,149***	1,143***	1,429***
<i>Cuidados Pessoais 2</i>	0,121***	0,122***	0,137***	0,138***	0,689***	0,859***
<i>Atividades Habituais 2</i>	0,095***	0,088***	0,108***	0,109***	0,503***	0,658***
<i>Dor/Desconforto 2</i>	0,067***	0,082***	0,079***	0,082***	0,436***	0,583***
<i>Ansiedade/Depressão 2</i>	0,062***	0,070***	0,066***	0,065***	0,405***	0,523***
<i>Mobilidade 3</i>	0,404***	0,404***	0,259***	0,266***	2,307***	3,107***
<i>Cuidados Pessoais 3</i>	0,247***	0,271***	0,171***	0,178***	1,052***	1,443***
<i>Atividades Habituais 3</i>	0,205***	0,228***	0,149***	0,162***	0,869***	1,286***
<i>Dor/Desconforto 3</i>	0,200***	0,244***	0,118***	0,127***	0,752***	1,078***
<i>Ansiedade/Depressão 3</i>	0,113***	0,135***	0,087***	0,099***	0,662***	0,922***
<i>Erro Absoluto Médio</i>	0,035	0,038	0,041	0,044	0,073	0,081
<i>No. &gt;0.05</i>	25	29	31	32	64	64
<i>No. &gt;0.10</i>	2	4	7	10	29	38
<i>rho de Pearson</i>	0,987	0,987	0,961	0,957	0,974	0,973
<i>rho de Spearman</i>	0,987	0,988	0,964	0,963	0,973	0,975

Fonte: elaboração dos autores.

\*\*\* Significante a 1%.

#### 4 DISCUSSÃO E CONCLUSÃO

Os resultados encontrados para o caso brasileiro apresentam uma taxa de inconsistência factível, dado que, quando se considera o conceito de monotonicidade estrita nas escolhas, apenas 8,7% dos pares comparáveis violam a ordem lógica esperada no TTO. Apesar da proporção de respondentes inconsistentes ser similar aquela achada na literatura internacional, as taxas de inconsistência foram significativamente superiores às encontradas por Badia, Monserrat e Herdman (1999), em estudo para a

Espanha que também utilizou o sistema descritivo EQ-5D, indicando a possibilidade de maior ocorrência de inconsistências no contexto de populações mais heterogêneas do ponto de vista socioeconômico (CRAIG e RAMACHANDRAN, 2006).

Esta evidência deve ser interpretada com cautela, na medida em que dois efeitos de sentidos opostos podem estar afetando os resultados. Primeiro, o desenho do estudo brasileiro diferiu de forma significativa de grande parte dos estudos internacionais. A utilização de blocos fixos com apenas oito estados de saúde, além do cartão morte, reduziu o número de pares de estados de saúde expostos a um mesmo indivíduo, se comparado à maioria dos estudos europeus que apresentam, em média, 13 estados de saúde por bloco. Dado o processo de aprendizagem adquirido ao longo da realização das tarefas (learning by doing), a maior taxa de inconsistência nas avaliações brasileiras podem resultar da reduzida exposição dos respondentes às avaliações. Segundo, a adoção de blocos com um menor número de estados de saúde podem se traduzir em menor nível de cansaço e morosidade na realização das atividades, resultando em menores níveis de inconsistência.

O Exercício de Troca de Tempo produziu maiores taxas de inconsistência que as demais métricas, independentemente da definição adotada na construção da taxa. Este resultado indica a existência de inconsistências secundárias, que são derivadas, em grande parte, por algum aspecto do procedimento de medida. Talvez não seja surpreendente que os métodos baseados na troca, como o time trade-off, produzam maior número de inconsistências que os demais, na medida em que as métricas de ordenação e EAV possuem a vantagem dos respondentes poderem visualizar todos os estados de saúde de forma simultânea, avaliando cada um deles com base não somente nos pontos de âncora, mas também com referência aos demais estados de saúde. Este não é o caso do exercício Time Trade-off, no qual os únicos pontos de referência são a Morte e a Saúde Perfeita (BADIA, MONSERRAT e HERDMAN, 1999). A maior frequência de valores idênticos para distintas e comparáveis condições de saúde, observada nos métodos cardinais, talvez possa ser explicada pela engenhosidade da métrica TTO e, em menor medida, da EAV. As consequências advindas desta complexidade no que se refere à compreensão dos exercícios por parte dos respondentes, ao alto grau de abstração necessário para a sua completude e à perda progressiva de estímulo, resultante da morosidade do instrumento, podem contribuir para a maior indiferença dos respondentes entre os estados de saúde passíveis de comparação (CRAIG e RAMACHANDRAN, 2006).

Segundo Devlin, Hansen e Selai (2004), a valoração de estados de saúde pode ser considerada um procedimento complexo de introspecção, avaliação e comparação, fazendo com que os respondentes lidem com situações hipotéticas, muitas vezes nunca experimentadas, e utilizem todas as funções cognitivas. Cada um dos métodos destaca, em seus procedimentos, diferentes processos cognitivos, fazendo com que diferenças consideráveis nos resultados sejam observadas (NORD, 1992). Assim, as particularidades de cada métrica, juntamente com o perfil do próprio respondente, impõem consequências importantes no que se refere à representatividade e à qualidade das informações. Esta evidência também deve ser considerada pelos pesquisadores no momento da escolha das métricas de mensuração de preferências.

O nível educacional e a condição socioeconômica foram determinantes importantes na geração de pares inconsistentes. A relação inversa entre nível de escolaridade e geração de inconsistências é extremamente forte no método Time Trade-off, de modo que qualquer ascensão no grau de escolaridade possui reflexos significativos na queda da inconsistência nesta métrica. Talvez, este resultado seja um reflexo do alto nível de abstração e peso cognitivo envolvidos na realização das tarefas referentes ao TTO. Apesar de significativo, o efeito da idade sobre a geração de inconsistências neste método foi irrisório. Quanto pior as condições materiais do domicílio, maiores foram as chances de se cometer ordenações de forma a violar a ordem lógica, em todos os métodos. Estes resultados mostram evidências da presença de inconsistências primárias, ou seja, àquelas geradas por limitações dos respondentes.

A qualidade dos dados depende da disposição do indivíduo em participar e revelar suas preferências de forma honesta e com suficiente nível de esforço (CRAIG e RAMACHANDRAN, 2006). Além disso, requer que o objeto em análise seja passível de classificação em uma escala de ordenação e que o participante tenha alguma capacidade cognitiva para realizar as tarefas, percebendo a relação entre suas decisões e o objetivo do estudo. A tarefa de avaliar estados de saúde pode ser cognitivamente difícil. A complexidade e a morosidade de determinados métodos podem imprimir cansaço ou confusão mental, principalmente aos idosos e aqueles com menor nível de escolaridade e condição socioeconômica. Dessa forma, alguns indivíduos podem falhar em entender as instruções, em compreender a natureza hipotética dos estados de saúde e ter dificuldades com o conceito de morte, contribuindo para o aumento de exercícios sem avaliações, questionários com apenas um ou dois estados avaliados ou com todos os estados apresentando o mesmo valor (CRAIG e RAMACHANDRAN, 2006).

De fato, quando se observa o perfil educacional e socioeconômico da população dos países europeus, onde foram pensadas e elaboradas as metodologias de avaliação de estados de saúde, as dificuldades intrínsecas aos exercícios não recebem destaque nas análises. Contudo, diante da expansão dos gastos privados e públicos com saúde experimentada por um número cada vez maior de países e a evidente importância da realização e incorporação de análises de custo-utilidade na alocação de recursos públicos, abre-se espaço para um debate mais crítico com respeito à utilização destes métodos em sociedades mais heterogêneas e com menor nível de desenvolvimento socioeconômico. A heterogeneidade cultural e socioeconômica se faz fortemente presente na América Latina, principalmente no Brasil. Enquanto no Canadá, Estados Unidos e Alemanha somente 2,4%, 1,7% e 5,0% da população com idade superior a 15 anos possuem apenas o ensino primário completo, no Chile, Argentina e Brasil, estas estimativas chegam a 12,3%, 32,7% e 27,3% em 2010 (THE WORLD BANK, 2013). Em termos de anos médios de estudo, o Brasil apresenta uma média de 7,5 anos de educação formal, seguindo em desvantagem em relação a alguns países latino-americanos (10,2 para Chile e 9,3 para Argentina) e desenvolvidos (12,1, 11,8 e 13,1 para Canadá, Alemanha e Estados Unidos, respectivamente). Além das discrepâncias internacionais, existem grandes disparidades internas ao país, que se refletem na desigualdade regional, no padrão dual de desenvolvimento socioeconômico representado pelos eixos Centro-Sul e Norte-Nordeste e entre classes econômicas. Neste sentido, questões como grau de abstração e peso cognitivo impostos pelo exercício, dificuldades de administração dos protocolos e custo de pesquisa devem se fazer presentes na escolha das métricas utilizadas.

Uma possível fragilidade deste trabalho refere-se ao conceito de inconsistências. A dificuldade intrínseca a esta conceitualização surge da complexidade de se julgar como inconsistentes, os comportamentos que violam a lógica do exercício praticado, quando, de fato, estes podem estar refletindo as preferências individuais. Uma evidência que ilustra este fato são os respondentes que experimentaram alguma condição de saúde ou trauma social e incorporaram a adaptação em suas avaliações, atenuando as diferenças percebidas entre estados de saúde (GROOT, 2000). Nesse contexto, pessoas idosas ou menos saudáveis possuem maior experiência com situações precárias de saúde e podem, via processo de adaptação, reverterem a ordem lógica esperada ou atribuírem valores semelhantes a estados de saúde marcadamente distintos no que se refere ao grau de severidade de seus domínios.

Em todos os métodos, a subamostra de indivíduos com, no máximo, uma inconsistência, apresentou o conjunto de pesos mais similares ao da amostra completa. Se considerarmos a censura de dados inconsistentes, os achados encontrados neste trabalho revelam que, para o caso de Minas Gerais, a exclusão destas informações afetaria muito sutilmente os pesos populacionais dos estados de saúde, em conformidade com os achados de Lamers *et al.* (2006), para o caso holandês. Talvez, dois fatores tenham contribuído para este resultado. Primeiro, a grande proporção de indivíduos consistentes ou com apenas uma inconsistência na amostra completa. Segundo, que a estrutura de preferência dos

excluídos talvez não difira muito daquelas de outros subgrupos ou da média populacional. Existem evidências de que, para Minas Gerais, enquanto indivíduos de grupos etários diferentes possuem utilidades estatisticamente distintas entre si, as variáveis de escolaridade e condição socioeconômica não são significativas para explicar diferenças das preferências médias (MESQUITA, 2013). Quando esta informação é aliada ao efeito ínfimo encontrado para a variável idade na explicação das inconsistências, pode-se inferir o baixo impacto da censura da amostra sobre os resultados encontrados.

Diferentemente deste estudo, altas proporções de respondentes com muitas inconsistências em suas avaliações podem resultar em subrepresentação na estimação dos pesos populacionais. Sociedades com alto grau de heterogeneidade socioeconômica e cultural ou menor nível de desenvolvimento econômico, associadas a diferenças intergrupos das estruturas de preferências, podem apresentar maior sensibilidade das tarifas em relação à censura dos dados inconsistentes. Em determinados países, algumas características sociodemográficas e de saúde têm mostrado possuir efeitos significativos sobre as preferências por estados de saúde (DOLAN *et al.*, 1996; KIND *et al.*, 1998; LUNDENBERG *et al.*, 1999). Neste sentido, parecem essenciais as tentativas de se melhorar a consistência por parte dos indivíduos idosos, com menores níveis de educação formal e de condição socioeconômica mais precária. Assim, a não ser que se invista em melhorias dos instrumentos de coleta de dados e no aperfeiçoamento destas técnicas a contextos sociais e demográficos distintos, existirá um desafio no que se refere à qualidade e à representatividade dos dados amostrais (DEVLIN *et al.*, 2003). Entretanto, avanços no controle de respostas inconsistentes a fim de se minimizar estes problemas podem resultar em informações que não refletem as decisões do respondente e, conseqüentemente, os gostos sociais. Neste sentido, uma possibilidade seria a ampliação do escopo de métricas de obtenção de pesos através de instrumentos alternativos. O menor nível de inconsistências encontrado no Exercício de Ordenação sugere suas potencialidades em avaliações econômicas em saúde.

## REFERÊNCIAS

ANDRADE, M. *et al.* Societal Preferences for EQ-5D Health States from a Brazilian Population Survey. *Value in Health Regional Issues*, v. 2, n. 3, p. 405-412, Dec. 2013.

ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA - ABEP. *Critério de classificação econômica Brasil*. Disponível em: <<http://www.abep.org/novo/Content.aspx?ContentID=835>>. Acesso em: 4 de fevereiro de 2013.

BADIA, X.; MONSERRAT, R.; HERDMAN, M. Inconsistent responses in three preference-elicitation methods for health states. *Social Science & Medicine*, v. 49, n. 7, p. 943-950, Oct. 1999.

BANSBACK, N. *et al.* Canadian Valuation of EQ-5D Health States: Preliminary Value Set and Considerations for Future Valuation Studies. *PLoS ONE*, v. 7, n. 2, p. 1-11, Feb. 2012.

BRAVATA, D. *et al.* Invariance and Inconsistency in Utility Ratings. *Medical Decision Making*, v. 25, n. 2, p. 158-167, 2005.

CHUANG, L.; KIND, P. The Effect of Health State Selection on the Valuation of EQ-5D. *Medical Decision Making*, v. 31, n. 1, p. 186-194, 2010.

CRAIG, B; BUSSCHABACH, J; SALOMON, J. Keep it simple: Ranking health states yields values similar to cardinal measurement approaches. *Journal of Clinical Epidemiology*, v. 62, p. 296-305, 2009a.

- CRAIG, B.; RAMACHANDRAN, S. Relative risk of a shuffled deck: a generalizable logical consistency criterion for sample selection in health state valuation studies. *Health Economics*, v. 15, n. 8, p. 835-848, Aug. 2006.
- DEVLIN, N.; HANSEN, P.; SELAI, C. Understanding health state valuation: A Qualitative analysis of respondents' comments. *Quality of Life Research*, v. 13, p. 1265-1277, 2004.
- DEVLIN, N. *et al.* Logical inconsistencies in survey respondents' health state valuations – a methodological challenge for estimating social tariffs. *Health Economics*, v. 12, n. 7, Jul. 2003.
- DOLAN, P. *et al.* The Time Trade-Off Method: results from a general population study. *Health Economics*, vol. 5, p. 141-154, 1996.
- DOLAN, P.; KIND, P. Inconsistency and Health State Valuations. *Soc. Sci. Med.*, v. 42, n. 4, p. 609-615, 1996.
- GROOT, W. Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics*, v. 19, n. 3, p. 403-420, 2000.
- GUDEX, C. *Time Trade-Off User Manual: Props and Self-Completion Method*. Centre for Health Economics, University of York, Occasional Paper Series, 1994.
- KIND, P. *Revised Protocol for the Valuation of Health States Defined by the EQ-5D-3L Classification System: Learning the Lessons from the MVH Study*. York: Centre for Health Economics, University of York, Nov. 2009.
- KIND, P. *et al.* Variations in population health status: results from a United Kingdom national questionnaire survey. *BMC Health Services Research*, v. 316, n. 7133, p. 736-41, Mar. 1998.
- KRABBE, P.; ESSINK-BOT, M.; BONSEL, G. The Comparability and Reliability of Five Health-State Valuation Methods. *Soc. Sci. Med.*, v. 45, n. 11, p. 1641-1652, 1997.
- LAMERS, L. The Transformation of Utilities for Health States Worse Than Death: Consequences for the Estimation for EQ-5D Value Sets. *Medical Care*, v. 45, n. 3, p. 238-244, Mar. 2007.
- LAMERS, L. *et al.* Inconsistencies in TTO and VAS Values for EQ-5D Health States. *Medical Decision Making*, v. 26, n. 2, p. 173-181, Mar-Apr. 2006.
- LEE, H. *et al.* Estimating weights for EQ-5D (EuroQol-5 dimensions) health states with the time trade-off method in Taiwan. *Journal of Formosan Medical Association*, v. 112, n. 11, p. 699-706, Nov. 2013.
- LUNDENBERG, L. *et al.* Health-states utilities in a general population in relation to age, gender and socioeconomic factors. *European Journal of Public Health*, v. 9, n. 3, p. 211-217, Mar. 1999.
- McCABE, C. *et al.* Using ranking data to estimate health state utility models. *Journal of Health Economics*, v. 25, n. 3, p. 418-431, Feb. 2006.

MESQUITA, K. *Heterogeneidade Individual e o Processo de Valorização dos Estados de Saúde em Minas Gerais*. Belo Horizonte: Centro de Desenvolvimento e Planejamento Regional, UFMG, 2013. Mimeografado.

NORD, E. Methods for Quality Adjustment of Life Years. *Soc. Sci. Med.* v. 34, n. 5, p. 559-569, 1992.

PHILLIPS, C. *What is a QALY?*. Health Economics. What is...? Series, 2<sup>nd</sup> ed. UK: Hayward, 2009.

SALOMON, J. Reconsidering the use of rankings in the valuation of health states: a model for estimating cardinal values from ordinal data. *Population Health Metrics*, v. 1, p. 1-12, Dec. 2003.

SIEGEL, S.; CASTELLAN Jr., N. *Estatística Não-Paramétrica Para Ciências do Comportamento*. 2 ed. Porto Alegre: Artmed, 2006. 448 p.

THE WORLD BANK GROUP. World Data Bank. Education Statistics, 2013. Disponível em: <<http://databank.worldbank.org/data/>>. Acesso em: 23/05/2013.

YUSOF, F.; GOH, A.; AZMI, S. Estimating an EQ-5D Value Set for Malaysia Using Time Trade-Off and Visual Analogue Methods. *Value in Health*, v. 15, n. Suppl, p. S85-S90, Jan-Feb. 2012.