

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL

ANDRÉ JACOB MORI

**Análise de inconsistências das preferências sociais
por saúde no Brasil através da utilização do
instrumento EQ-5D**

Belo Horizonte
2016

ANDRÉ JACOB MORI

**Análise de inconsistências das preferências sociais por
saúde no Brasil através da utilização do instrumento
EQ-5D**

Dissertação de Mestrado apresentada ao Departamento de Pós-Graduação em Economia do Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Economia.

Orientadora: Prof^a. Dr^a. Mônica Viegas Andrade
Coorientadora: Prof^a. Dr^a. Kenya Noronha

Belo Horizonte
2016

M854a 2016	<p>Mori, André Jacob. Análise de inconsistências das preferências sociais por saúde no Brasil através da utilização do instrumento EQ-5D [manuscrito] / André Jacob Mori. - 2016 118 f., enc. : il.</p> <p>Orientadora: Mônica Viegas Andrade. Coorientadora: Kenya Noronha. Dissertação (mestrado). Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. Inclui bibliografia (f. 102-108) e anexos.</p> <p>1. Saúde - Aspectos econômicos - Teses. 2. Economia - Teses. I. Andrade, Mônica Viegas. II. Noronha, Kenya. III. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. IV. Título.</p> <p>CDD: 338.473621</p>
---------------	--

AGRADECIMENTOS

Agradeço a minha família por sempre ter apoiado e incentivado, dando condições para estudar e buscar o crescimento através da dedicação e o empenho. Dedico especiais agradecimentos às minhas orientadoras Mônica Viegas Andrade e Kenya Noronha pelo apoio, ajuda e paciência que tiveram comigo desde o início desta empreitada. Dedicaram dias, noites e até madrugadas junto a mim para que este trabalho fosse possível.

Agradeço imensamente a minha esposa Léa pelo apoio, amor e carinho incondicionais ao longo de toda esta jornada, desde ler diversas vezes o texto até dando todo o apoio nos momentos mais difíceis. Ao longo deste trabalho pudemos desfrutar dos melhores momentos de nossas vidas.

Agradeço a todos os professores responsáveis pela minha formação: professores e funcionários do Cedeplar e da USP. Por todo o apoio e os incentivos que oferecem aos alunos e pela confiança que depositam nos mestrandos em todas as nossas atividades.

Agradeço também aos amigos Lucas e Carla, que trabalham excepcionalmente bem, além de ajudar e possibilitar que este trabalho possa ser entregue.

RESUMO

A estimação das preferências sociais por estados de saúde através do instrumento EQ-5D já foi realizada em diversos países, em sua maioria implementando o protocolo MVH Original. Este trabalho tem como objetivos comparar duas pesquisas realizadas no Brasil que utilizaram protocolos de aplicação similares, mas com desenhos de estudo baseados em blocos fixos no caso da pesquisa realizada em Minas Gerais e da coleta de informações para todos os estados de saúde oriundos do EQ-5D na pesquisa realizada em 3 capitais de estado. A análise será baseada na presença de inconsistências lógicas cometidas pelos entrevistados. A análise leva em conta a natureza e a descrição dos estados de saúde, como também os potenciais efeitos das características individuais sobre a consistência lógica dos respondentes. Foram utilizadas duas bases de dados, coletadas entre 2011 e 2012, uma no estado de Minas Gerais e outra nas cidades de Porto Alegre, Recife e Rio de Janeiro. Na pesquisa de campo são investigados indivíduos alfabetizados, com idade entre 18 e 64 anos, residentes de áreas urbanas. A partir de alguns algoritmos matemáticos calculam-se as inconsistências lógicas em termos individuais e também em termos da distância entre os pares lógicos comparáveis. A comparação entre as duas pesquisas se baseou em três métodos de obtenção das preferências, ranking, escala visual analógica e *Time Trade Off*. Um dos resultados encontrados é que as inconsistências lógicas se concentram em estados de saúde próximos, indicando que a maior parte destes desvios decorre da dificuldade em perceber a diferença na descrição dos estados de saúde. A pesquisa realizada nas capitais apresentou taxas de inconsistências mais elevadas, indicando que a presença de estados de saúde não balanceados pode ter influenciado a capacidade de compreensão de alguns indivíduos. Também foi proposta a exclusão seletiva de indivíduos que cometeram desvios lógicos entre estados de saúde muito distantes, no caso da pesquisa de Minas Gerais a qualidade do ajuste dos modelos de estimação dos pesos sociais por estados de saúde melhora com a exclusão, mas este efeito não pode ser verificado na pesquisa das capitais. Os resultados apresentados nesta dissertação sugerem a viabilidade do uso do desenho de estudo que utiliza blocos fixos, que também podem reduzir os custos logísticos de implementação da pesquisa

Palavras-chave: Preferências sociais. Desenhos de estudo. Comparabilidade. Inconsistências lógicas. EQ-5D. Utilidade.

ABSTRACT

The elicitation of societal preferences through EQ-5D instrument has been implemented in several countries, in most cases based on the original MVH protocol developed by the EuroQol Group. This paper aims to compare two studies conducted in Brazil based on similar protocols that differed only in the study design. The first survey was conducted in the Minas Gerais state, and was based on a fixed balanced group of health states. The second study occurred in 3 different cities: Porto Alegre, Recife and Rio de Janeiro and collected social preferences for all of the 243 health states from EQ-5D. The analysis will be based in the presence of inconsistent health state valuations. This phenomena can arise both from the description and nature of health states as well as the individual characteristics that may affect the valuations. Both studies occurred between 2011 (Minas Gerais) and 2012 (Rio de Janeiro, Recife and Porto Alegre). The interviewers surveyed literate individuals between the ages of 18 and 64, living in urban areas. The logical inconsistencies were estimated through mathematical algorithms and were accounted in individual levels and in terms of distance of pairwise health state valuations. Three elicitation methods were analyzed: ranking, visual analog scale and Time Trade Off. The logical inconsistencies concentrate in health states that are close to each other, which can indicate that individuals fail to identify the differences in the health state descriptions. Higher inconsistency levels were found in the second survey, which might indicate that the balancing of health states is a desirable feature in terms of consistent responses. Another result showed that the selective exclusion of individuals based on the distance between pairwise inconsistent valuations can improve the goodness of fit of the estimated social tariffs models, both on visual analog scale and TTO. These results could not be reproduced in the unbalanced study design. These paper findings might suggest that the balanced fixed block study design is a desirable feature when estimating societal preferences for health state, which might also reduce logistical costs of conducting the field research.

Keywords: Valuation methods. Study Design. Comparability. Logical inconsistencies. EQ-5D. Utility.

LISTA DE FIGURAS

FIGURA 1. Representação Gráfica do <i>Standard Gamble</i>	28
FIGURA 2. Representação Gráfica do Time Trade Off	29
FIGURA 3. Histograma do total de pares de estados de saúde comparáveis por amostra	66
FIGURA 4. Distribuição dos pares de estados de saúde comparáveis por amostra	67
FIGURA 5. Proporção de Inconsistências Lógicas por distância para o Ranking, EVA e Time Trade Off, pesquisa das capitais e do Estado de Minas Gerais.....	88
FIGURA 6. Média ponderada da taxa de inconsistência para o ranking, EVA e TTO, pesquisa das capitais e do Estado de Minas Gerais.....	89
FIGURA 7. Comparação dos valores preditos para o TTO para a amostra total e para 2 sub-amostras- Pesquisa das capitais e do estado de MG	96
FIGURA 8. Comparação dos valores preditos para a EVA para a amostra total e para 2 sub-amostras- Pesquisa das capitais e Estado de Minas Gerais.....	97

LISTA DE TABELAS

Tabela 1. Sistema Descritivo EuroQol: dimensões de saúde e pontuações do EQ-5D-3L.....	22
Tabela 2. Estados de saúde avaliados no protocolo MVH.....	34
Tabela 3. Estados de Saúde utilizados no estudo norte-americano.....	36
Tabela 4. Estados de saúde utilizados no estudo dinamarquês	37
Tabela 5. Características sociodemográficas da amostra de Minas Gerais e comparação outras bases de dados.....	55
Tabela 6. Características Socioeconômicas Amostra Nacional.....	58
Tabela 7. Distribuição das distâncias entre estados de saúde e a saúde perfeita	61
Tabela 8. Estados de saúde utilizados na pesquisado Estado de Minas Gerais	62
Tabela 9. Variáveis utilizadas nos modelos Logísticos	73
Tabela 10. Descrição das variáveis utilizadas nos modelos de estimação dos pesos para estados de saúde.....	76
Tabela 11. Proporção e Taxas de Inconsistências individuais para ambas as amostras	80
Tabela 12. Teste para igualdade nas taxas de inconsistência média ponderadas.....	81
Tabela 13. Taxa média de inconsistência lógica por características sócio demográficas e de saúde	83
Tabela 14. Resultados para o Modelo Geral Logístico Multivariado – Ranking, EVA e TTO- Amostra das Capitais e de Minas Gerais	84
Tabela 15. Resultados para o Modelo Geral Logístico Multivariado – Ranking, EVA e TTO- Amostra das Capitais e de Minas Gerais	93
Tabela 16. Qualidade do Ajuste com Exclusão de Indivíduos Inconsistentes em cada Sub-Amostra, TTO, EVA: Amostra das Capitais.....	95
Tabela 17. Qualidade do Ajuste com Exclusão de Indivíduos Inconsistentes em cada Sub-Amostra, TTO, EVA: Amostra de Minas Gerais.....	95

Sumário

1.	Introdução	12
2.	Análises econômicas em saúde	15
2.2	EQ-5D-3L	20
2.3	Medidas cardinais de estimação da utilidade.....	23
2.3.1	Métodos baseados na teoria da escolha.....	25
	Fonte: Cruz (2010)	28
2.3.2	Escala analógica e exercício de ordenamento.....	31
2.4	O Protocolo original MVH	32
	Fonte: Gudex <i>et al.</i> 1994.....	34
2.4.1	Revisões sugeridas ao protocolo original.....	35
3	Inconsistências nas preferências por estados de saúde	39
3.1	Efeitos do método de aplicação das pesquisas.....	41
3.2	Evidências empíricas internacionais	43
3.3	Estudos aplicados no Brasil	47
4.	Metodologia	51
4.1	Bases de dados	51
4.1.1	Amostra e coleta de dados do estudo de Minas Gerais.....	52
4.1.2	Amostra do estudo das três cidades	56
4.2	Aplicação do protocolo da entrevista.....	58
4.2.1-	Diferenças no Desenho dos Estudos.....	60
4.3	Comparabilidade e Inconsistências Lógicas	64
4.3.1	Métricas de cálculo de inconsistências lógicas	68
4.4-	Análise das Inconsistências segundo características individuais.....	70
4.4.1	Modelo Logit.....	70

4.4.2 Variáveis utilizadas no modelo	72
4.5 – Estimação e comparação dos modelos de regressão para a EVA e TTO	74
5. Resultados	78
5.1. Métricas de inconsistências lógicas	78
5.2 Associação entre características Individuais e a taxa de inconsistência	81
5.3 Relação entre inconsistências lógicas e as distâncias dos pares de estados de saúde	88
5.3.1 Modelos logísticos considerando as distâncias:.....	90
5.3.2 Qualidade de ajuste de Modelos realizando exclusões seletivas baseadas na distância entre pares de estados de saúde.....	94
6. Considerações Finais.....	99
REFERÊNCIAS.....	104
ANEXO A - INSTRUMENTOS	111
ANEXO B – QUADRO DA TROCA DE TEMPO	112
FIGURA A – QUADRO DE TEMPO 1: Estados melhores que a morte	112
FIGURA B – QUADRO DE TEMPO 2: Estados piores que a morte.....	112
PROTOCOLO DA TROCA DE TEMPO	118
SEÇÃO I – ESTADO MELHOR QUE A MORTE.....	119
SEÇÃO II – ESTADO PIOR QUE A MORTE.....	120

1. Introdução

Informações técnicas necessárias para a tomada de decisão de políticas públicas de saúde frequentemente estão indisponíveis ou não são apresentadas na forma correta. A avaliação de tecnologias é uma forma de pesquisa que identifica os problemas, avalia os impactos de intervenções alternativas e apresenta suas conclusões. Os tomadores de decisão perceberam que a avaliação de tecnologias em saúde é cada vez mais importante, especialmente no contexto de rápido desenvolvimento tecnológico verificado desde o início do século XX até os dias atuais (BATTISTA, HODGE, 1999). A utilização das demandas sociais como subsídio para a avaliação de tecnologias tem ganhado relevância e a obtenção das preferências sociais se faz cada vez mais necessária.

Dentro do escopo das análises econômicas em saúde, as medidas que levam em conta os custos monetários e avaliam os benefícios em termos de preferências sociais são conhecidas como QALY (*Quality Adjusted Life Years*) ou AVAQ (Anos de vida ajustados à qualidade). Esta medida combina as dimensões de mortalidade e também de qualidade de vida relacionada a saúde (PHILIPS, 2009). Williams (1985) descreve os AVAQ como uma estrutura flexível que engloba uma ampla variedade de processos de avaliação numa medida simples e versátil de sucesso. O AVAQ leva em conta uma possível troca entre qualidade de vida por longevidade, a qual estabelece um *trade-off*.

Estas medidas que incorporam a troca de qualidade de vida por quantidade (tempo) são bastante utilizadas em análises econômicas, especialmente pelo fato de serem fundamentadas no conceito microeconômico de utilidade esperada. Os exercícios que permitem a coleta das preferências individuais que incorporam esse *trade-off* costumam ser cognitivamente desafiadores, com necessidade de abstração por parte dos indivíduos e aplicação complexa para os entrevistadores, sendo necessária a adoção de hipóteses teóricas implícitas na sua realização (LAMERS, 2007; ROBISON, LOOMES e JONES-LEE, 2001; CRAIG, 2009).

O processo de valoração dos estados de saúde é complexo, uma vez que envolve diversas capacidades cognitivas, como avaliação, comparação e abstração. Os indivíduos precisam se

imaginar em situações de saúde diferentes da sua, comparar e refletir sobre a descrição destes estados de saúde e também se atentar para a distância relativa entre eles (DOLAN, 2000). Muitas vezes, essas situações ou estados de saúde são hipotéticos e nunca antes vivenciados, o que envolve, para sua valoração, habilidades de percepção, memória, aprendizado e expressão. Devlin, Hansen e Selai (2004) apontam que este processo é desafiador e pouco usual à maioria dos indivíduos, Nord (1992) destaca que cada exercício propõe um desafio diferente ao entrevistado, que juntamente com suas próprias características individuais acaba produzindo resultados distintos para cada atividade. A qualidade e representatividade das informações deve levar em conta estas características captadas em cada forma de obtenção das preferências individuais.

Os métodos de estimação e coleta das preferências sociais por saúde foram inicialmente propostos e construídos em países desenvolvidos, cujo perfil socioeconômico é muito diferente daquele encontrado na população brasileira. Nas avaliações realizadas nestes países, o nível educacional verificado é muito superior àquele encontrado no Brasil. Além disso, o país também apresenta uma grande desigualdade regional, com eixo desenvolvido concentrado nas regiões Sul e Sudeste e polos de elevada desigualdade de renda e concentração de pobreza no Norte-Nordeste. Duas pesquisas já foram conduzidas e publicadas para o Brasil utilizando o sistema descritivo do EQ-5D, uma no estado de Minas Gerais (ANDRADE *et al.*, 2013) e uma pesquisa que amplia os resultados de Minas Gerais, incorporando também as capitais Porto Alegre, Recife e Rio de Janeiro (SANTOS *et al.*, 2015). A primeira pesquisa de campo foi conduzida ao longo do ano de 2011, enquanto a segunda em 2012.

Estes trabalhos propõem a estimação das preferências sociais por estados de saúde no país, permitindo a comparação com os demais estudos produzidos ao redor do mundo. Aqui será investigada a presença de avaliações inconsistentes entre os pares de estados de saúde produzidas nas pesquisas realizadas no país. Será enfatizada a diferença metodológica observada os dois estudos, separando os resultados para a pesquisa realizada no estado de Minas Gerais e a pesquisa realizada nas capitais.

Os resultados verificados neste trabalho apontam que a utilização de blocos de estados de saúde balanceados, proposta na pesquisa realizada em Minas Gerais, produz resultados satisfatórios nos exercícios do *Time Trade Off* e também na escala visual analógica em relação à pesquisa realizada nas capitais. A diferença metodológica na coleta das informações para os estados de saúde requer uma logística mais custosa em termos de pesquisa de campo, uma vez que o EQ-5D produz 243 estados de saúde diferentes.

Além disso, verifica-se também que a maior parte destas inconsistências se concentra em estados de saúde com descrições muito parecidas, isto é, com distâncias muito próximas entre si. Esta análise já foi proposta em trabalhos anteriores (BADIA, HERDMANN e MONSERAT, 1999), mas com amostras menores e sem o alcance nacional.

Este trabalho está dividido em cinco seções, o capítulo 2 contextualiza e motiva a discussão das análises econômicas em saúde, com ênfase na análise de custo utilidade. O capítulo 3 avança na análise de inconsistências lógicas, explorando os trabalhos já realizados em outros países para subsidiar a discussão e as metodologias implementadas neste trabalho. No capítulo 4 de metodologias, serão descritos os algoritmos implementados para os cálculos de inconsistências lógicas e os modelos utilizados para estimar as taxas de inconsistências médias ponderadas e por distância, nesta seção será explicitada a diferença metodológica existente entre as pesquisas de Minas Gerais e a pesquisa das capitais. O capítulo 5 contém os resultados obtidos empregando os modelos previamente apresentados e no capítulo 6 é realizada a comparação entre os resultados verificados nas pesquisas de Minas Gerais e das capitais, com ênfase especial no diferente desenho de estudo utilizado nas duas pesquisas e seus efeitos sobre as inconsistências lógicas.

2. Análises econômicas em saúde

A avaliação de tecnologias em saúde tem sido um importante instrumento de análise para a decisão de incorporação de inovações tecnológicas no setor de saúde. Desde o pós-guerra, as pesquisas em saúde se multiplicaram e novas tecnologias e métricas de avaliação estão sendo desenvolvidas e implementadas em muitos países (WARNER; LUCE, 1982).

A Avaliação de Tecnologias em Saúde (ATS) pode ser conceituada como um processo contínuo de avaliação de tecnologias, que visa ao estudo sistemático de suas consequências. Seu objetivo é prover informações para a tomada de decisão, tanto política quanto clínica em relação a incorporação e descarte de tecnologias. A avaliação de tecnologias em saúde inclui tanto os aspectos relacionados aos benefícios ou resultados obtidos com determinada tecnologia como também a dimensão dos custos. Desse modo, a grande vantagem da ATS é propiciar para o tomador de decisão um parâmetro comparável associado a cada tecnologia que permite, na medida do possível, abarcar as propriedades associadas aos resultados e aos custos. A propriedade essencial da ATS é fornecer ferramentas aos gestores para a tomada de decisões econômicas sobre a comparação entre tratamentos alternativos, com caráter multidisciplinar e abrangente (BRASIL, 2009).

No Brasil, o acesso público aos serviços de saúde é oferecido através do SUS (Sistema Único de Saúde), estabelecido na constituição (BRASIL, 1988). Uma das políticas norteadoras do sistema é a equidade no acesso e garantia de universalidade. Segundo Carvalho e Santos (1995), as diretrizes básicas do sistema são descentralização, atendimento integral, participação da comunidade e a equidade.

Como existe escassez de recursos voltados para a área da saúde, a garantia de assistência integral à população representa um desafio para o Sistema de Saúde (BRASIL, 2009). O sistema privado de serviços coexiste no país e essa estrutura dual oferece ainda mais desafios para as decisões sobre políticas públicas. O Estado deve agir tanto como provedor de recursos públicos, quanto como regulador dos serviços prestados pelos agentes privados. Dentro desta perspectiva, a ATS será um importante instrumento para ajudar na escolha econômica da alocação de recursos.

No Brasil, a discussão se inicia na década de 80, mas apenas em 2003, com esforços conjuntos do Ministério da Ciência e Tecnologia e o Ministério da Saúde, foram implementadas iniciativas importantes para o desenvolvimento e incorporação de avaliações de tecnologias em saúde. Em 2004, foi criado o Departamento de Ciência e Tecnologia (DECIT), responsável pela avaliação de tecnologias em saúde dentro do SUS. No ano de 2008 foi criada a REBRATS (Rede Brasileira de Avaliação de Tecnologias em Saúde), cujo fim é oferecer pareceres técnico-científicos ao governo na tomada de decisão sobre políticas públicas alternativas.

Os países desenvolvidos como Austrália, Estados Unidos, Canadá e Inglaterra já implementam análises econômicas em saúde há vários anos, através de agências nacionais que subsidiam a escolha entre intervenções públicas alternativas (BRASIL, 2008). Os EUA têm destaque no desenvolvimento de estudos neste sentido, em especial para tratamentos e intervenções médicas específicas (BANTA, LUCE; 1983). Estudos mais recente mostram que nos EUA a utilização de avaliação de tecnologias tem sido explicitamente incorporada à política de cobertura oficial do país (Medicare) e para a definição dos serviços de saúde oferecidos aos funcionários públicos (Blue Shift e Blue Cross) no qual é obrigatória a existência de estudos que comprovem a eficácia, eficiência e segurança dos tratamentos oferecidos à população (KRAUS-SILVA, 2003; TUNIS E KANG, 2001)

Na Austrália, as políticas de avaliação em saúde foram implementadas de forma sistemática a partir de 1982, com duas instituições responsáveis pela avaliação de intervenções médicas (*Medical Services Advisory Committee- MSAC*) e em medicamentos (*Pharmaceutical Benefit Advisory Committee- PBAC*). Estas instituições priorizam suas análises em três aspectos fundamentais: segurança, efetividade e custo-efetividade (JACKSON, 2007 e NEUMANN, 2005).

Na Inglaterra, o *National Institute for Health and Clinical Excellence (NICE)*, que é a agência Nacional de avaliação de políticas em saúde, tem uma missão mais abrangente do que a simples recomendação de intervenções na área da saúde, sendo também responsável pelo estabelecimento de critérios de padronização das análises com o objetivo de homogeneizar as avaliações. A

definição dos segmentos, setores e insumos voltados para a área da saúde é guiada pelos estudos realizados pelo NICE (NEUMANN, 2005; BRASIL, 2008).

No Canadá, a *Canadian Agency for Drugs and Technology in Health* (CADTH) foi criada em 1989 sob o nome de *Canadian Coordinating Office for Health Technology Assessment* -CCOHTA (HAILEY, 2007), este órgão tem atuação relevante na orientação aos governos estaduais e provinciais do país, responsáveis pela prestação dos serviços básicos de saúde. A agência tem atuação nacional e atua em conjunto com as universidades e administrações públicas na implementação de políticas de saúde mais eficientes. Um exemplo da integração entre a universidade e a atuação das agências reguladoras canadenses foi o desenvolvimento de probióticos voltados para a prevenção e tratamento para diarreia, realizado pela universidade de McGill em 2005.

Os países em desenvolvimento começaram a promover as análises econômicas para a escolha de tecnologias de forma mais eficiente, especialmente com a menor disponibilidade de recursos existente na maioria deles (SECOLI *et al.*, 2005). Segundo Médici (1994), os gastos com saúde desembolsados pelos países na América Latina variaram entre 2% a 9% do PIB ao longo da década de 1990. Segundo Costa, Barbosa e Zucchi (2015), os gastos com saúde, considerando união, estados e municípios somam 4,02% do PIB em 2013. Dados da OMS apontam que países desenvolvidos gastaram cerca de 10% do PIB no setor de saúde, como Canadá (10,9%), França (11,7%), Alemanha (11,3%), Reino Unido (9,1%) e Estados Unidos (17,1%) (WHO, 2014).

Os estudos econômicos em saúde mais utilizados se baseiam nas análises de custo-minimização, custo-benefício, custo-efetividade ou custo-utilidade. A característica comum destas avaliações é que o custo é sempre medido em termos monetários e o que as difere é a unidade de medida do benefício. Na análise de custo-minimização, em geral as intervenções propiciam um mesmo benefício e a escolha é feita com base no menor custo de implementação. A restrição desta abordagem é que não se pode comparar intervenções que produzam desfechos diferentes. Na análise custo-benefício os desfechos da intervenção são calculados em termos monetários, o que pode gerar críticas, especialmente em questões éticas sobre como atribuir valor à vida humana.

Estipular um valor monetário a um benefício de algum tratamento também pode ser uma tarefa complexa e arbitrária (HURLEY, 2010).

A análise custo-efetividade mede os benefícios em unidades dos resultados do desfecho de uma intervenção, como por exemplo número de internações evitadas e anos de vida salvos. A intervenção em estudo definirá a unidade de benefício analisada. A principal vantagem desta métrica é sua fácil compreensão, podendo ser empregada por todos os tipos de gestores de políticas de saúde, como médicos, enfermeiros e demais profissionais da área de saúde. Sua desvantagem é que não permite avaliar intervenções que tenham efeitos em mais de uma dimensão (redução nas internações e também no período de internação, por exemplo), além de também não permitir a comparação entre intervenções que levem a diferentes resultados (HURLEY, 2010; GOLD, FRANKS e ERICKSON, 1996; DRUMMOND *et al.*, 1997).

A avaliação de custo-utilidade difere das outras duas abordagens ao medir a unidade de benefício através de ganhos de utilidade aos pacientes sendo usual a medida de Anos de Vida ajustados a Qualidade (AVAQ). Essa métrica tem a vantagem de permitir a comparação entre políticas alternativas, considerando tanto a dimensão de mortalidade quanto a morbidade. A morbidade dos pacientes é aferida através de um instrumento que analisa a qualidade de vida, incluindo diferentes dimensões do estado de saúde (TORRANCE; FEENY, 1989).

A análise custo-utilidade é mais indicada quando a qualidade de vida é o resultado mais relevante. Nesses casos, é importante considerar como consequência da intervenção tanto as modificações na mortalidade como na morbidade. Por utilizar como medida de resultado os anos de vida ajustados a qualidade (AVAQ), a abordagem de custo-utilidade tem a propriedade de incorporar essas duas dimensões. Uma vantagem desta abordagem é a possibilidade de se comparar diferentes estratégias de intervenção em saúde direcionadas a diferentes condições de saúde, como por exemplo avaliações do custo-utilidade de uma estratégia de tratamento de câncer que pode ser comparada a uma estratégia de tratamento de dislipidemia, considerando o benefício em saúde da população como AVAQs (Brasil, 2008).

O uso deste tipo de análise econômica não é recomendável quando os resultados em termos de saúde não são claros ou disponíveis, quando as intervenções analisadas têm resultados igualmente efetivos ou quando a qualidade de vida é importante, mas pode ser obtida através de uma única variável, com unidades naturais, ou quando o custo de obtenção e utilização das utilidades é elevado ao ponto de inviabilizar a análise ou incapaz de alterar os resultados do estudo (TORRANCE, 1986). Torrance (1986) afirma que a implementação deste tipo de análise costuma ter custos elevados, mas que conforme mais estudos são publicados e utilizados, maior a capacidade de comparação entre eles, tornando esta abordagem cada vez mais prática.

Do ponto de vista teórico, a construção do AVAQ supõe que os indivíduos auferem utilidade considerando o tempo de sobrevivência e a qualidade de vida. A hipótese principal é que os indivíduos estão dispostos a trocar quantidade (tempo) por qualidade de vida. Tratamentos que elevam as chances de sobrevivência impondo restrições às condições de saúde podem gerar menos AVAQ do que um tratamento menos restritivo. Por exemplo, um ano de vida adicional sob saúde perfeita corresponde a 1 AVAQ; caso o estado de saúde seja debilitado e avaliado com utilidade de 0,5, que corresponde a um decréscimo de 50% em relação à saúde perfeita, seriam necessários 2 anos a mais de expectativa de vida para que os dois tratamentos fossem equiparados em termos de AVAQ.

O conceito de AVAQ foi utilizado pela primeira vez por Klarman *et al.* (1968) sendo construído pelo número de anos de vida ponderado por um índice ou peso relacionado à qualidade de vida em cada ano. O autor comparou pacientes renais crônicos tratados com hemodiálise e transplante de rim. Verificou que um ano de vida realizando hemodiálise correspondia a 75% da qualidade de vida de um paciente transplantado.

O uso do AVAQ é indicado quando as intervenções afetam as dimensões de longevidade e da qualidade de vida. Alguns estudos apontam que tratamentos que elevam a expectativa de vida, normalmente serão acompanhados por algum tipo de restrição ou debilidade (ROMERO, LEITE e SWARCSWALD, 2005; CAMARGOS, PERPÉTUO e MACHADO, 2005). Um país que tem especial destaque no uso dos AVAQs na avaliação de tecnologias é o Reino Unido, através do NICE. No contexto britânico, a medida do AVAQ tem especial relevância já que a população do

país tem elevada expectativa de vida, 81,5 anos de idade segundo dados do *Office for National Statistics* – ONS UK (2012), a população de idosos é crescente e doenças crônicas tem uma maior prevalência sobre estes indivíduos.

O envelhecimento da população global está prestes a atingir um novo patamar, o número de idosos vai superar pela primeira vez na história o número de crianças nos próximos cinco anos, segundo a OMS (WHO, 2011). Esse novo perfil demográfico terá impactos importantes sobre a formulação de políticas de saúde, com ênfase nas doenças crônicas, que afetam mais os indivíduos de idade avançada. Métricas que levem em conta a qualidade de vida deverão ganhar espaço, e este papel é bem desempenhado pelos AVAQs. O tratamento de doenças crônicas como Alzheimer, diabetes e Parkinson, é caro e necessita de tecnologias e terapias mais avançadas, a problemática principal destas doenças não é sua letalidade, mas a restrição gradativa das atividades cotidianas, reduzindo a qualidade de vida dos indivíduos (WHO, 2011).

Em países em desenvolvimento, como o Brasil, coexistem perfis epidemiológicos muito distintos, parte da população ainda é impactada por doenças relacionadas à pobreza, enquanto outras regiões apresentam maiores problemas em doenças crônicas e degenerativas. As patologias que afligem as populações mais pobres (como doenças infectocontagiosas, parasitárias ou decorrentes de subnutrição) estão mais relacionadas com a mortalidade do que com o bem-estar. As enfermidades crônicas e degenerativas têm maiores impactos nas populações mais envelhecidas e afetam mais a qualidade de vida dos pacientes. Nesta realidade dual presente no Brasil, as preferências sociais por estados de saúde apresentarão um perfil distinto daquele observado em países desenvolvidos. A utilização dos anos de vida ajustados à qualidade é especialmente importante no país pelo processo de envelhecimento da população, já que os indicadores de saúde tradicionais não são capazes de captar a dimensão de qualidade de vida.

2.2 EQ-5D-3L

O cálculo dos AVAQs associado a diferentes eventos em saúde deve considerar o tempo de sobrevivência dos indivíduos condicionado a ocorrência do evento e a qualidade de vida. A principal dificuldade diz respeito à mensuração da qualidade de vida, uma vez que é necessário não só avaliar o estado de saúde dos indivíduos após determinado evento, como mensurar a perda de

utilidade associada a esses estados de saúde que sucedem determinado evento. A avaliação da qualidade requer o uso de instrumento que descreva o estado de saúde dos indivíduos. Diversos sistemas descritivos foram desenvolvidos para descrever os estados de saúde.

Não existe um sistema que seja capaz de descrever de uma forma suficiente todas as dimensões dos estados de saúde humanos (dimensões físicas, psicológicas e emocionais). Dessa forma existe uma gama de instrumentos que procuram contemplar as diferentes dimensões do estado de saúde. Esses instrumentos podem ser subdivididos entre genéricos e específicos. Os instrumentos genéricos possibilitam uma análise direta, podem ser utilizados para descrever o estado de saúde de grupos populacionais distintos e cobrem diversas formas de morbidade ou patologias. Já os instrumentos específicos são desenhados para avaliar condições particulares de vida, geralmente indicados para pacientes que apresentam doenças crônicas, problemas em funções (função sexual, por exemplo) ou grupos populacionais determinados (idosos, mulheres, crianças). Estes instrumentos são capazes de captar com maior sensibilidade clínica os domínios de relevância do aspecto a ser avaliado (CAMPOLINA e CICONELLI, 2006).

Dentro da multiplicidade de sistemas de classificação implementados para a construção dos AVAQ, destacam-se o *Health Utility Index (HUI)*, o SF-36, sua versão reduzida SF-6D e o EQ-5D. O HUI, que compreende três sistemas descritivos HUI-1, HUI-2 e HUI-3, é exigente em termos amostrais, e descreve estados de saúde através de 8 dimensões (GOLD, FRANKS e ERICKSON, 1996; FEENY *et al* 2002). O SF-36 e sua versão reduzida SF-6D (*Short-Form 36* e *Short Form 6 dimensions*, respectivamente) descrevem os estados de saúde dentro de 8 categorias possíveis: funcionamento físico, limitações devido aos problemas físicos, funcionamento social, dor corporal, limitações devido a problemas emocionais, saúde mental, vitalidade e percepção geral da saúde. (BRAZIER *et al.*, 1992; BRAZIER, ROBERTS e DEVERILL, 2002; HURST, RUTA e KIND, 1998). Cruz (2011) realizou pesquisas na população brasileira implementando o SF-6D para a obtenção das preferências sociais por saúde na cidade de Porto Alegre, no Rio Grande do Sul.

O instrumento EuroQol 5 dimensões (EQ-5D-3L) foi desenvolvido pelo grupo de pesquisadores do EuroQol, ao longo da década de 1990. Um grupo de pesquisadores que utilizou o sistema

descritivo para realizar o estudo pioneiro MVH (*Measurement and Valuations of Health Group*), utilizado pelo sistema de saúde britânico na avaliação de políticas de saúde NICE (*National Institute for Health and Clinical Excellence*) (NICE, 2008). É o instrumento mais difundido entre os apresentados por sua simplicidade e maior abrangência.

A formulação do conteúdo descritivo presente no EQ-5D partiu da revisão de instrumentos que mediam a qualidade de vida relacionada à saúde, do inglês HRQoL, sintetizando de maneira mais direta os estados de saúde (GUDEX *et al.*, 1994). A ideia do questionário original era, devido a sua simplicidade, possibilitar que o instrumento fosse preenchido através de pesquisas postais (BUSSCHBACH *et al.*, 2005). O EQ-5D descreve o estado de saúde individual através de um conjunto de 5 dimensões: mobilidade, autocuidado, atividades usuais, dor/desconforto e ansiedade ou depressão. Cada dimensão apresenta 3 níveis de gravidade: nenhum problema (nível 1), presença de algum problema (nível 2), e presença de problema grave (nível 3). A tabela abaixo apresenta o instrumento EQ-5D-3L traduzido para o português e validado para a sua utilização considerando a população brasileira.

Tabela 1. Sistema Descritivo EuroQol: dimensões de saúde e pontuações do EQ-5D-3L

Dimensão	Nível	Pontuação
Mobilidade	Não tenho problemas para andar	1
	Tenho alguns problemas para andar	2
	Estou limitado a ficar na cama	3
Cuidados Pessoais	Não tenho problemas com meus cuidados pessoais	1
	Tenho alguns problemas para me lavar ou me vestir	2
	Sou incapaz de me lavar ou de me vestir sozinho	3
Atividade Habituais	Não tenho problemas para desempenhar minhas atividades usuais	1
	Tenho alguns problemas para desempenhar minhas atividades habituais	2
	Sou incapaz de desempenhar minhas atividades habituais	3
Dor/Desconforto	Não tenho dores ou mal-estar	1
	Tenho dores ou mal-estar moderado	2
	Tenho dores ou mal-estar extremos	3
Ansiedade/Depressão	Não estou ansioso(a) ou deprimido(a)	1
	Estou moderadamente ansioso(a) ou deprimido(a)	2
	Estou extremamente ansioso(a) ou deprimido(a)	3

Fonte: Adaptação para o português de Dolan e Roberts (2002)

A combinação de 5 dimensões com três níveis de gravidade resulta na definição de 243 estados de saúde distintos. Esses estados são representados por uma combinação de cinco dígitos que variam de 1 a 3 cada. Os estados extremos são descritos como 11111, saúde perfeita, ou seja, nenhum problema em todas as 5 dimensões, e 33333, equivalente ao pior estado de saúde, com problemas extremos em todas as dimensões. Um exemplo de estado de saúde descrito pelo EQ-5D é o estado 12312, em que o indivíduo não tem problemas para andar (1), tem alguns problemas para se lavar ou se vestir (2), é incapaz de desempenhar suas atividades habituais (3), não tem dores ou mal-estar (1) e é moderadamente ansioso ou deprimido (2).

Mais recentemente foi desenvolvido o instrumento EQ-5D-5L no qual cada uma das cinco dimensões de saúde apresenta cinco níveis de gravidade. Estes níveis de gravidade diferem-se dos três níveis verificados no EQ-5D-3L introduzindo a ideia de maior gradação. Existirão os seguintes níveis: (1) não tenho problemas de locomoção; (2) tenho problemas leves de locomoção; (3) tenho problemas de locomoção moderados; (4) tenho graves problemas de locomoção, (5) sou incapaz de me locomover, tradução própria, (HERDMAN *et al.* 2011, tradução própria). A ideia era melhor distinguir as condições de saúde moderadas. No entanto, o seu uso ainda é limitado e não há ainda um consenso sobre os ganhos obtidos dessa versão em relação ao EQ-5D-3L (GOLICKI *et al.* 2015). Estes autores afirmam ainda que o maior espectro de níveis de estados de saúde dificulta a implementação de algoritmos capazes de estimar os pesos para estados de saúde, que podem ser diretamente estimados na modalidade com 3 níveis de gravidade.

2.3 Medidas cardinais de estimação da utilidade

A construção dos AVAQS requer, além da definição de um sistema descritivo, que sejam estimadas as preferências dos indivíduos pelos 243 estados de saúde definidos pelo EQ-5D-3L. Uma vez determinados estes pesos, poder-se-á então comparar eventos ou tecnologias alternativas em termos de anos de vida ajustados à qualidade ganhos com cada uma das possíveis intervenções. A intervenção que oferecer um maior nível de utilidade a um custo em unidades

monetárias mais baixo deverá ser escolhida em detrimento das demais intervenções ou tecnologias. Por exemplo, se uma intervenção clínica possibilitar o ganho de 1 AVAQ ao custo de 1000 unidades monetárias e uma alternativa oferecer ganho de 0,8 AVAQ ao custo de 1000 unidades monetárias, então a primeira tecnologia deverá ser escolhida pelo formulador de política pública ou gestor de saúde.

A estimação das preferências sociais por estados de saúde deve ocorrer, preferencialmente através de pesquisas que considerem a população de cada país, uma vez que as preferências dos indivíduos sobre cada estado de saúde dependem do contexto socioeconômico, cultural e da forma de organização e provimento dos serviços de saúde.

Como visto, a estimação dos pesos necessários para a construção dos AVAQs é um processo extenso e complexo, de modo que a utilização desta métrica para avaliação de tecnologias deve se beneficiar de seus diferenciais em relação às demais análises econômicas, a variação na qualidade de vida decorrente de cada evento ou tecnologia deverá ser um aspecto relevante para a tomada de decisão. A implementação desta metodologia é indicada quando os resultados das intervenções afetam diferentes dimensões de saúde e a qualidade de vida é uma dimensão de interesse.

Os países pioneiros que realizaram pesquisas nacionais para estimação de preferências por estados de saúde foram Inglaterra (GUDEX *et al.* 1994), Holanda (KRABBE, ESSINK-BOT e BONSEL; 1997), Espanha (BADIA; MONSERAT; HERDMAN, 1999). Norman *et al.* (2009) realizam uma comparação entre as pesquisas nacionais em que o EQ-5D foi implementado. Os autores verificaram que existem diferenças significativas entre as valorações dos estados de saúde entre 10 estudos realizados em diferentes países. Os autores reforçam a importância de se aplicar pesquisas nacionais visando a obtenção das preferências sociais por estados de saúde. Neste sentido, mais países também vem implementando estes estudos nacionais, como Estados Unidos (SHAW, COONS E JOHNSON, 2002), Coreia do Sul (LEE *et al.*, 2012), França (CHEVALIER E POURVOURVILLE, 2013). Mais recentemente, países em desenvolvimento começaram a realizar estes estudos. Na América Latina essas pesquisas ocorreram na Argentina

(AUGUSTOVSKI, ARES e GIBBONS, 2009), Chile(ZARATE, V. *et al*; 2011) e Brasil (ANDRADE *et al.*, 2013).

A estimação das preferências sociais pode ser realizada a partir de diferentes métodos. Os dois métodos mais comuns empregados para os estudos de valoração dos estados de saúde empregam o *Standard Gamble*(SG) e o *Time Trade Off* (TTO), que são baseados na teoria da escolha. Os métodos da Escala Analógica Visual (EVA) e o Ranking usualmente são utilizados como preparação e familiarização dos respondentes das pesquisas, mas recentemente alguns estudos têm utilizado esses métodos mais simples para estimar as preferências sociais por estados de saúde

2.3.1 Métodos baseados na teoria da escolha

O *Standard Gamble* é um método baseado na Teoria da Utilidade Esperada, de Von Neuman Morgenstern (1944),no qual os indivíduos precisam fazer escolhas diante de situações sob incerteza. A função de utilidade representa as preferências individuais diante de uma loteria. Esta loteria configura uma espécie de aposta, que consiste em dois cenários com probabilidade de ocorrência diferentes. Torrance (1986) justifica que a teoria de utilidade esperada é adequada para a maioria dos bensquantificáveis e o uso desta teoria se disseminou também na área da saúde (CRUZ, 2010).

Para que essa função de utilidade seja capaz de representar preferências,estas deverão ser racionais e respeitar o axioma da independência. Preferências racionais requerem que as hipóteses de preferências completas e transitividade sejam respeitadas. Preferências completas pressupõe que o indivíduo seja capaz de comparar todos os bens disponíveis em um determinado espaço de escolha. No contexto de preferências por estados de saúde, o conceito de relação de preferências completa prevê que o indivíduo possa se imaginar vivendo hipoteticamente nos estados de saúde por um determinado período de tempo, e então revelar sua escolha. O indivíduo deve ser capaz de classificar os estados de saúde como preferíveis ou indiferentes entre si.

A hipótese de transitividade é atendida quando a relação de escolha entre alternativas respeita uma hierarquia. Considerando três desfechos, A, B e C, se A é preferido a B; B é preferido a C,

então A deverá ser preferido a C, isto é, deverá ser capaz de comparar alternativas e não violar sua hierarquia de escolha. Ao respeitar a premissa de transitividade, é possível inferir as preferências do indivíduo com relação às várias alternativas disponíveis sem ser necessário realizar todas as comparações possíveis. No contexto das preferências por estados de saúde, isso implica que os indivíduos devem ser capazes de estabelecer uma determinada consistência ao comparar estados. Situações de saúde mais debilitadas devem gerar menor bem-estar ao indivíduo e devem, portanto, receber menores valorações.

O axioma da independência prevê que a relação entre dois desfechos deve se preservar, com a introdução de novas alternativas. Se A é preferível a B, A+C deverá ser preferível a B+C. No contexto de obtenção de preferências por saúde, o nível de utilidade atribuído a determinado estado de saúde não deverá sofrer influências dos demais eventos da loteria (Brazier *et al.*, 1999).

No método *Standard Gamble* o indivíduo realiza uma escolha entre resultados alternativos, sendo que em um deles existe um componente de incerteza. Este tipo de exercício é indicado para pesquisas em saúde pelo componente de incerteza presente na análise, nenhum tratamento médico está isento de riscos, assim como nenhum tratamento é sempre eficaz em todas as intervenções.

São apresentadas duas alternativas: opção A refere-se a um tratamento com dois resultados possíveis: com probabilidade p , o paciente retorna à saúde perfeita e permanece neste estado por t anos adicionais; ou o paciente morre de forma imediata com probabilidade $1-p$. Assim, representando a função de utilidade individual como $V = p.h_i$, onde h_i refere-se a condição de saúde i (neste caso será um estado de saúde hipotético oriundo do sistema descritivo), a utilidade esperada nesta aposta corresponde a $V_A = p.1 + (1-p).0 = p$. A opção B apresenta um resultado certo: o indivíduo vive por t anos em uma determinada condição de saúde. A utilidade esperada nesta aposta corresponde a $V_B = 1.h_i = h_i$. A probabilidade p sofre alterações até que o respondente seja indiferente entre as duas alternativas.

O ponto de indiferença do indivíduo é atingido pela igualdade entre as utilidades derivadas nas duas alternativas, ou seja, $V_A = V_B \leftrightarrow h_i = p$. Logo, a utilidade do estado de saúde avaliado é igual

a probabilidade p . A probabilidade de sucesso associada à indiferença do respondente entre duas alternativas provê uma unidade de medida para o valor do estado de saúde. Se o estado de saúde for muito indesejável, o indivíduo aceita uma intervenção para alterar sua condição mesmo com uma baixa probabilidade de recuperar sua saúde.

Para estados considerados piores que a morte, a alternativa sob incerteza oferece ao indivíduo um resultado de saúde perfeita com probabilidade p ou uma determinada condição de saúde com probabilidade $1-p$, enquanto a alternativa sob certeza refere-se à morte. Assim, as utilidades associadas às alternativas são $V_A = p \cdot 1 + (1-p) \cdot h_i$ e $V_B = 1 \cdot 0 = 0$.

Neste caso a utilidade do estado de saúde avaliado h_i é igual a $-p/(1-p)$. Se um estado de saúde é extremamente indesejável, um indivíduo provavelmente não escolheria a aposta, a não ser que a probabilidade de retornar ao estado de saúde perfeita fosse alta. Estados piores que a morte são representados por números negativos. O método sempre envolve uma escolha entre um desfecho de certeza e outro de incerteza, onde o estado de saúde associado ao desfecho de ocorrência certa é intermediário em termos de preferência em relação ao melhor (11111) e pior estado (33333). Como a técnica baseia-se na tomada de decisão sob um ambiente de incerteza, respeitando o referencial teórico da Teoria da Utilidade Esperada, o uso do *Standard Gamble* é considerado ideal em análises econômicas em saúde. A figura 1 representa o *Standard Gamble*. A ideia do *trade-off* longevidade e qualidade de vida pode-se verificar claramente neste exercício

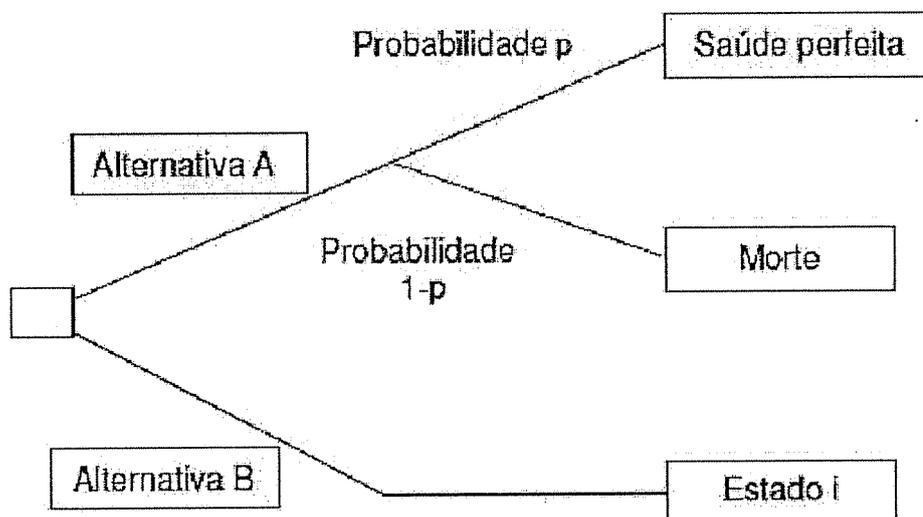


FIGURA 1. Representação Gráfica do *Standard Gamble*

Fonte: Cruz (2010)

Apesar do SG ser um exercício indicado para estimar as preferências sociais por estados de saúde, a aplicação deste método em populações com baixo grau de escolaridade pode levar a problemas de compreensão dos indivíduos, especialmente em relação aos cenários hipotéticos de saúde. Apesar de a hipótese de probabilidade ser um ponto central no que se refere a intervenções clínicas, como probabilidade de sucesso ou falha, este conceito pode ser complexo para alguns indivíduos. Caso a compreensão do conceito de probabilidade não seja adequada, a coleta das informações de parte significativa da população pode ser comprometida. O *Standard Gamble* é um exercício probabilístico em essência.

Thomas, Sacklet e Torrance (1972) desenvolveram o *Time Trade Off* (TTO) buscando diminuir o problema imposto pelo conceito de probabilidade. Neste método, são propostas duas situações de certeza ao indivíduo, que envolvem troca de quantidade e qualidade de vida num ambiente de escolha determinístico. Primeiramente o indivíduo deve escolher se o estado de saúde apresentado é melhor ou pior do que a morte imediata, após esta escolha, o exercício segue por uma das etapas descritas a seguir.

São oferecidas aos indivíduos duas alternativas. Na primeira, o indivíduo vive t anos sob alguma condição de saúde seguidos pela morte, e na segunda alternativa o indivíduo vive $x < t$ anos com saúde perfeita, também seguidos pela morte. O entrevistador deve deixar x variar até que se verifique indiferença entre alternativas. O tempo de exposição dos indivíduos ao estado de saúde debilitado tem efeitos sobre a percepção deles sobre esta condição de saúde. Dolan (1996) aponta que estados de saúde mais graves tornam mais intoleráveis à medida que o tempo passa. Nas pesquisas implementadas na maioria dos países, adotou-se como padrão a medida $t=10$ anos.

Quanto mais próximo da saúde perfeita for o estado avaliado, menos tempo um sujeito estaria disposto a trocar entre quantidade e qualidade de vida. Indivíduos que não toleram perda de qualidade de vida tenderão a trocar mais anos de saúde perfeita por anos sob condições de morbidade. A função de utilidade será definida como $V = Th_i$, onde T são anos de vida, e V o valor da utilidade, que será a fração em anos de vida em saúde perfeita equivalente em anos de vida sob algum estado de saúde, $h_i = \frac{x}{t}$. A figura 2 retrata graficamente a escolha do indivíduo sob o TTO.

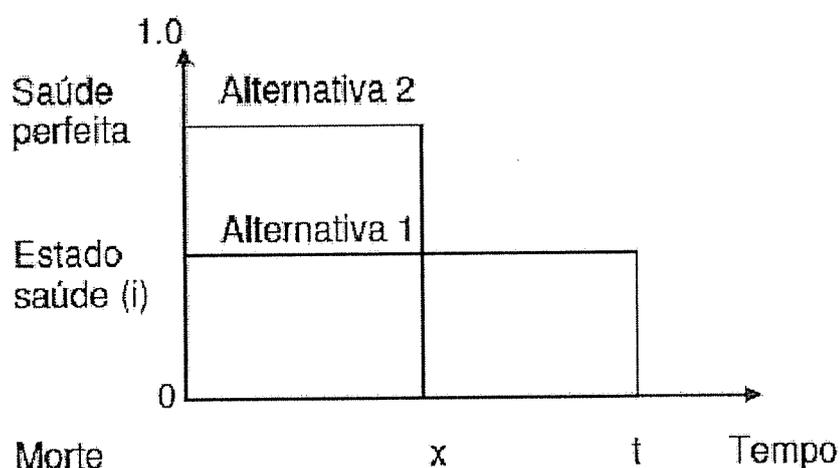


FIGURA 2. Representação Gráfica do Time Trade Off

Fonte: Cruz (2010)

O exercício de troca de tempo também permite que o estado avaliado seja considerado pior do que a morte. A troca será entre estar morto e a alternativa de viver x anos na condição de saúde

considerada pior do que a morte, seguida por $(t - x)$ anos sob saúde perfeita. A duração de x deve variar até que a indiferença seja alcançada. Neste caso o valor de utilidade será calculado como

$$-\frac{x}{t-x}$$

Tanto o *Time Trade Off* quanto o *Standard Gamble* são exercícios que partem do princípio de *trade-off* entre longevidade e qualidade, características desejáveis para realização de estudos econômicos em saúde. O TTO identifica diferentes pontos da curva de indiferença individual, sob o arcabouço da teoria da utilidade esperada e sob um ambiente de escolha determinística. O exercício ainda envolve trocas de componentes da função de utilidade individual, e oferece uma alternativa factível ao *Standard Gamble* em termos conceituais e de aplicação prática. Métodos baseados na teoria da escolha são mais indicados para obtenção das preferências por estados de saúde pois permitem que indivíduos realizem trocas, de modo que se torna possível observar como ocorre a variação do bem-estar diante de mudanças das condições de saúde.

A desvantagem da abordagem baseada na escolha é o nível de abstração e habilidade cognitivas necessárias para realização dos exercícios (COAST *et al.*, 2008). A violação dos axiomas da teoria da utilidade esperada pode levantar questionamentos se estas métricas podem ou não ser implementadas em determinados contextos (BRAZIER *et al.*, 1999).

O estudo de Dolan (1996) verificou que alguns estados de saúde se tornam gradativamente intoleráveis conforme o tempo em que o indivíduo está exposto a determinada condição de saúde. Mostrou que o efeito do tempo sobre percepção do estado de saúde não responde de forma linear e que análises de sensibilidade são necessárias para captar estes efeitos. A janela fixa de tempo em que o respondente revela suas preferências entre estados alternativos normalmente empregada nos estudos é de 10 anos.

Loomes e McKenzie (1989) também relatam que a atitude em relação ao risco do entrevistado pode se alterar diante de certas condições de saúde. Uma hipótese defendida pelos autores é que a percepção de riscos do agente econômico depende do tempo em que está exposto a determinado evento e também sofre alterações com o ciclo de vida do indivíduo. Um indivíduo jovem tende a atribuir um menor peso à saúde perfeita do que um adulto que acaba de se aposentar, isso sugere

que os indivíduos tendem a atribuir maior valor à saúde em certos estágios de suas vidas. Uma medida linear de peso para o estado de saúde pode ser incapaz de captar essa característica.

Kahneman e Tversky (1974) também apontam evidências que a percepção de entrevistados pode se alterar de acordo com a forma de enquadramento da questão. Durante a aplicação do protocolo, as respostas do indivíduo podem ser afetadas caso os estados de saúde apresentados possuam diferenças muito pequenas entre si ou diferenças substanciais. A escolha dos estados de saúde apresentados ao indivíduo pode induzir respostas diferentes dentro de uma mesma metodologia.

A hipótese base do TTO é que os indivíduos estão dispostos a trocar uma quantidade constante de anos de vida por qualidade de vida, independentemente da idade em que se encontram. Dependendo do tempo de exposição a determinada condição de saúde, os indivíduos podem se adaptar ou então se tornar intolerantes, alterando suas preferências pelos estados de saúde e consequentemente violando o axioma da independência (BRAZIER, 1999).

2.3.2 Escala analógica e exercício de ordenamento

Os exercícios de classificação são implementados como preparação para as metodologias mais complexas como o TTO e o SG. O intuito de realizar estes exercícios é familiarizar os indivíduos à descrição dos estados de saúde, oriunda do sistema descritivo. Recentemente estas medidas tem ganhado relevância para a estimação das preferências por estados de saúde pela sua simplicidade e fácil obtenção (CRAIG, SOLOMON e BUSSBACH; 2009).

A escala analógica visual é um exercício mais simples, em que o indivíduo deve classificar estados de saúde alternativos dentro de uma escala padronizada, frequentemente associada a um termômetro, com valores de 0 a 100 (KRABBE *et al.*, 2006). As extremidades desta escala em geral são categorizadas como saúde perfeita e morte. Pode-se produzir uma versão variante em que o limite inferior da escala representa o pior estado de saúde produzido pelo sistema descritivo permitindo assim que os estados de saúde sejam classificados como piores do que a morte.

Pela simplicidade e fácil aplicação deste exercício, alguns autores defendem sua utilização para a construção dos parâmetros de valorização dos estados de saúde, em especial em populações com menor grau de escolaridade (BANSBACK, 2012). A menor exigência cognitiva necessária para classificar estados na escala permite a aplicação deste exercício para populações com menor grau de instrução. Em geral, a estimação dos pesos sociais por estados de saúde é feita pela EVA nas pesquisas postais que utilizam o EQ-5D (DOLAN, KIND, 1996; OHINMAA, SINTONEN, 1999; DEVLIN, KIND, 2000).

Existem críticas sobre esse método para aferir preferências sociais por saúde, por não se basear em teorias de escolha (ROBISON, LOOMES, JONES-LEE, 2001; KRABBE, ESSINK-BOT, BONSEL 1997). No entanto, estudos mostram que as valorações estimadas através de modelos Logit condicionais EVA utilizando os resultados do ordenamento possuem correlação elevada com os resultados estimados com base no TTO, desde que sejam feitos ajustes adequados em sua mensuração (CRAIG, SOLOMON e BUSSCHBACH; 2009).

O exercício do ordenamento é o mais simples de todos os apresentados, após receber os cartões que contém a descrição dos estados de saúde, o indivíduo deve ordená-los de forma crescente em relação às suas preferências. Coletar informações baseando-se no ranking também pode ser uma alternativa viável para implementação da pesquisa, reduzindo substancialmente os gastos com treinamento de entrevistadores, uma vez que uma descrição simples de como proceder com a execução do exercício deve ser suficiente para qualquer indivíduo alfabetizado. Craig *et al.* (2009) sugere que até mesmo crianças podem realizar o ordenamento, como um jogo recreativo de ordenamento. Brazier *et al.* (2012) alertam, no entanto, que as informações coletadas sob o ordenamento e o EVA não apresentam ancoragem em algum estado. O conceito de trade-off entre qualidade de vida e longevidade não está evidenciado nestas duas abordagens,

2.40 Protocolo original MVH

O protocolo original MVH (GUDEX, 1994) foi desenvolvido e utilizado no estudo pioneiro do EQ-5D-3L, realizado na Inglaterra em 1994. O protocolo de aplicação desta pesquisa foi

replicado na maior parte dos outros estudos, em alguns casos com adaptações, mas preservando a mesma estrutura.

Devido à dificuldade de se avaliar diretamente os 243 estados de saúde, o protocolo MVH define a seleção de 45 estados de saúde descritos pelo EQ-5D (42 estados de saúde, a morte; a saúde perfeita e o estado inconsciente) para serem diretamente valorados. Esses estados são classificados em subconjuntos de 13 estados de saúde que deveriam ser avaliados pelos indivíduos. A coleta de dados para os 243 estados originados no EQ-5D é muito exigente em termos amostrais. Dessa forma, no protocolo original, 43 estados de saúde eram investigados para a construção dos parâmetros de valorização dos estados de saúde, e os pesos para os demais estados eram estimados através de modelos de regressão. Estudos apontaram que os indivíduos não eram capazes de avaliar mais do que 13 estados de saúde (LAMERS *et al.* 2006), principalmente devido ao longo tempo de aplicação das entrevistas. Além disso, o sistema descritivo é capaz de produzir alguns estados de saúde que podem ser pouco factíveis na realidade, como por exemplo o estado 31111, no qual o indivíduo está limitado a ficar na cama, mas não tem problemas em realizar seus cuidados pessoais e tarefas cotidianas, não apresenta dor ou desconforto, não é ansioso ou depressivo.

Os 43 estados de saúde utilizados no protocolo MVH estavam amplamente distribuídos no espaço de valoração, para que incluíssem todas as possíveis combinações entre as dimensões e os graus de gravidade gerados pelo sistema descritivo, possibilitando a existência de interações entre os estados. Os estados de saúde utilizados na pesquisa estão na tabela 1 abaixo. Todos os indivíduos avaliaram os estados âncora, receberam 2 dos 5 estados mais leve, 3 dos 12 estados leves, 3 dos 12 estados moderados e finalmente 3 dos 12 estados graves.

Tabela 2. Estados de saúde avaliados no protocolo MVH

Estados âncora	11111; 33333; inconsciente
Estados mais brandos	11112; 11121; 11211; 12111; 21111
Estados brandos	11122; 11131; 11133; 21133; 21222; 21312; 12211; 11133; 22121; 12121; 22112; 11312
Estados Moderados	13212; 32331; 13311; 22122; 12222; 21323; 32211; 12223; 22331; 21232; 32313; 22222
Estados severos	33232; 23232; 23321; 13332; 22233; 22323; 32223; 32232; 33321; 33323; 23313; 33212

Fonte: Gudex *et al.* 1994

No protocolo do MVH, os entrevistadores solicitam primeiramente que os indivíduos reportem seu estado de saúde atual dentro das 5 dimensões de saúde do sistema descritivo e atribuam uma nota ao seu estado de saúde que varia de zero a 100 em uma escala visual analógica (EVA). A EVA é uma escala similar a um termômetro de 20 centímetros que varia entre zero e 100. A realização desta tarefa permite que os indivíduos visualizem as dimensões existentes no instrumento e a descrição dos níveis de gravidade de cada uma delas. Posteriormente, utilizam-se os exercícios de classificação (ordenamento) e da escala visual analógica para familiarizar o entrevistado com o sistema descritivo para então prosseguir para o exercício de troca do tempo (Time Trade Off - TTO). No exercício da ordenação, os indivíduos classificam os estados de saúde definidos pelo EQ-5D-3L, descritos em cartões, em ordem crescente, de acordo com suas preferências. O intuito de realizar este exercício é familiarizar os indivíduos com a descrição dos estados de saúde hipotéticos que serão valorados pelo método do TTO.

Logo em seguida, os indivíduos classificam os mesmos estados de saúde na EVA. No protocolo original, o indivíduo realizava a classificação dos estados na escala logo após o ordenamento.

A pesquisa prossegue então para a realização do *Time Trade Off*, através do quadro de Troca de Tempo, apresentado nos anexos deste trabalho. Este quadro apresenta duas faces, uma para estados considerados melhores do que a morte e outra para estados considerados piores do que a morte. Na face em que o estado é considerado melhor do que morte, o indivíduo deve apontar a quantidade x ($x < 10$) de anos em saúde perfeita que estaria disposto a trocar por 10 anos na condição de saúde apresentada pelo estado de saúde avaliado. Quanto menor o valor de x , pior é

o estado de saúde de acordo com a avaliação de cada entrevistado. Se o estado for considerado pior que a morte, os indivíduos escolhem entre a morte imediata ou então viver um número x de anos ($x < 10$) sob a condição de saúde descrita pelo estado, seguida por $(10-x)$ anos em saúde perfeita, como forma de compensação. Dessa forma, quanto maior a compensação necessária (quanto menor o x), pior é este estado.

2.4.1 Revisões sugeridas ao protocolo original

Embora não exista um consenso formal sobre as revisões do protocolo MVH, alguns pesquisadores verificaram a necessidade de alterações pontuais sobre o protocolo de aplicação da pesquisa. A revisão de certas características do protocolo se baseou em seus elementos, a primeira delas foi em relação à escolha dos estados de saúde utilizados na pesquisa. Estudos implementados nos EUA, Alemanha e Dinamarca (SHAW, COONS, JOHNSON; 2002; GREINER *et al*, 2005; WITTRUP-JENSEN, LAURIDSEN, PEDERSEN, 2008) utilizaram blocos fixos de estados de saúde.

Shaw, Coons e Johnson (2002) optaram por uma escolha de estados de saúde apresentada na tabela 3. Sua pesquisa visava a estimação das preferências sociais por saúde nos EUA, utilizaram os mesmos 45 estados do protocolo MVH, mas cada indivíduo deveria avaliar um grupo de 15 estados de saúde, ao invés dos 13 estados da pesquisa pioneira. Cada respondente atribuiu valores aos estados âncora, estado inconsciente e 33333, 2 estados muito leves e um conjunto aleatório de 9 estados de saúde oriundos da tabela 2. A tabela 3 apresenta os grupos de estados distribuídos aleatoriamente aos respondentes da pesquisa.

Tabela 3. Estados de Saúde utilizados no estudo norte-americano

Amostras para Modelagem				Amostra para validação
Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
11111	11111	11111	11111	11111
33333	33333	33333	33333	33333
Morte Imediata	Morte Imediata	Morte Imediata	Morte Imediata	Morte Imediata
Inconsciente	Inconsciente	Inconsciente	Inconsciente	Inconsciente
11121	12111	21111	11211	12111
11211	21111	11112	11112	11121
11131	11113	11122	22121	22121
21222	21133	12211	11133	11133
12121	21312	22112	11312	21222
13311	12222	32331	22122	22331
21323	32211	32313	13212	12222
12223	22331	22222	21232	13311
33232	13332	23321	23232	23321
32223	22323	22233	33323	33321
32232	33212	23313	33321	23313

FONTE: Shaw, Coons e Johnson 2012, tradução própria

Greiner *et al.* (2005) utilizaram 36 estados de saúde em sua pesquisa para estimar as preferências sociais por estados de saúde na Alemanha. Estes estados foram escolhidos dentre os perfis de saúde propostos pelo protocolo original, e os autores justificam que com este conjunto já é possível realizar as estimações de modelos extrapolando os resultados para os demais estados de saúde disponíveis no sistema descritivo, permitindo também uma amostra menos numerosa. Cada entrevistado atribuiu valores para conjuntos que variavam entre 13 e 15 estados de saúde.

Wittrup-Jensen, Lauridsen e Pedersen (2008) propõe também a escolha dos estados de saúde diferente daquela empregada pelo protocolo MVH. Os critérios empregados pelos autores para a escolha foram: todos os indivíduos deveriam avaliar os estados de saúde perfeita (11111), o pior estado de saúde (33333), o estado “morte” e também o estado 22222. Além destes, deveriam atribuir valores a 2 dos estados mais leves, na mesma categorização do protocolo original, 8 estados selecionados de forma a cobrir os estados de saúde mais sérios e mais leves, e finalmente 2 estados de saúde relevantes que pacientes com diabetes ou doenças cardíacas comumente definiam sua própria saúde. Dessa forma, cada respondente avaliava 16 estados de saúde. Os autores dividiram sua amostra em 4 sub-amostras que avaliaram os estados conforme a tabela 4 a seguir.

Tabela 4. Estados de saúde utilizados no estudo dinamarquês

Amostra 1	Amostra 2	Amostra 3	Amostra 4	Comentário
11111	11111	11111	11111	
22222	22222	22222	22222	
33333	33333	33333	33333	Os quatro estados comuns em nass 4
morte	morte	morte	morte	sub-amostras
21111	11112	11211	21111	
11211	12111	11121	11112	2 dos 5 estados mais brandos
12211	11131	23313	32232	
11113	22121	11122	11312	
21222	21312	13311	11133	
32331	12222	22122	33321	8 estados usados em outros estudos,
23321	32211	22331	21323	abrindo todos os níveis
21232	12223	22112	13212	
23232	21133	32313	Inconsciente	
22233	22323	32223	33232	
11212	21221	21121	11221	Dois estados comumente relatados
22333	33322	21322	22322	por pacientes de diabetes/cardiácos

FONTE: Wittrup-Jensen, Lauridsen e Pedersen (2008, tradução própria)

Kind (2009) sugere algumas alterações que foram implementadas em estudos mais recentes (CHEVALIER e POURVOURVILLE, 2013; LEE *et al.*, 2009), inclusive no estudo aplicado no estado de Minas Gerais (ANDRADE *et al.*, 2013). A escolha de estados utilizada no protocolo original (avaliação direta de 45 estados) reduzia o escopo de estados disponíveis para examinar as diferenças individuais (KIND, 2009). Uma sugestão para aumentar o número de estados de saúde valorizados em campo foi utilizar uma estrutura de blocos de estados de saúde, com a avaliação direta de 102 estados de saúde ao invés dos 45 originais. Esta modificação propõe uma categorização dos estados de saúde baseado na distância entre o estado e a saúde perfeita.

Baseando-se neste conceito de gravidade dos estados de saúde, os 102 estados foram subdivididos em 26 grupos, cada um contendo 2 estados leves, 2 estados moderados e 2 estados graves. A sugestão de escolha dos estados de saúde proposta nesta revisão e implementada na pesquisa realizada para o estado de Minas Gerais (ANDRADE *et al.* 2013) segue a lógica de blocos fixos de estados de saúde.

Outra modificação sugerida é que entre a execução do ordenamento e a realização do exercício na EVA recomenda-se que os estados de saúde sejam embaralhados. O intuito desta alteração é verificar se existe uma coerência do respondente ao analisar os estados com uma escala de valores e sem a presença da escala. Antes dessa modificação, como os cartões com a descrição dos estados de saúde não eram embaralhados, os indivíduos na maior parte das vezes preservavam a ordenação do exercício de ordenamento para atribuir os valores na EVA. Como resultado, a correlação entre os dois exercícios era próxima de 1 (CRAIG, BUSSCHBACH E SALOMON, 2009).

Uma modificação mais simples que também foi sugerida foi a eliminação do estado de saúde inconsciente, que não é descrito pelas dimensões disponíveis no sistema descritivo e que poderia ser omitido sem perda de explicação relevante para o protocolo. Kind (2009) também sugere mudar a forma como o estado morte era apresentado. O autor sugere que a utilização do termo “morte” pode levar alguns entrevistados a imaginar situações como eutanásia, por exemplo, e que isso possa ter implicações sobre o julgamento dos respondentes, mas mesmo assim o autor orienta que o termo seja substituído por “morte imediata”.

3 Inconsistências nas preferências por estados de saúde

No contexto das preferências por estados de saúde, um aspecto fundamental diz respeito à presença de avaliações inconsistentes. Como mencionado anteriormente, a teoria da utilidade esperada pressupõe racionalidade das preferências. Esse pressuposto garante consistência ao comportamento do consumidor em relação às suas escolhas. Uma escolha individual é definida inconsistente quando viola uma relação lógica de dominância entre estados de saúde. Para ilustrar este problema, tomemos os estados 11222 e 11333. Observa-se que o primeiro estado apresenta problemas moderados nas dimensões de atividades usuais, dor e desconforto e de ansiedade ou depressão, enquanto que o segundo estado apresenta problemas graves nas mesmas três dimensões. Como o segundo estado é igual ou pior em cada uma das dimensões, é logicamente inferior ao primeiro. Caso receba uma avaliação superior ao estado 11222, uma inconsistência lógica foi cometida.

Especificamente, em relação às preferências por estados de saúde, a presença de avaliações inconsistentes pode estar associada a diferentes aspectos. Um primeiro aspecto diz respeito ao entendimento do método utilizado para aferição das preferências. O exercício da troca de tempo, como explicado, é complexo e requer capacidade cognitiva dos indivíduos tanto para entender a descrição dos estados de saúde, como para realizar as trocas entre qualidade e quantidade de tempo de vida. No exercício da troca de tempo há ainda uma dificuldade que pode resultar em avaliações inconsistentes devido à sequencialidade envolvida na avaliação dos diferentes estados de saúde: os indivíduos valoram cada estado separadamente, podendo ocorrer avaliações inconsistentes em decorrência do “efeito memória” (ANDRADE *et al.*, 2016).

Um segundo aspecto que explica a presença de inconsistências diz respeito a qualquer processo de escolha que envolve diversas opções. Os indivíduos podem ser inconsistentes, sobretudo por violar transitividade, por diversos motivos intrínsecos a esse processo. A forma ou a ordem como as opções são apresentadas, por exemplo, pode resultar em escolhas inconsistentes; indivíduos podem mudar suas preferências ao longo do processo de avaliação tanto por um aprendizado com essas escolhas como por depender das escolhas passadas (ANDRADE *et al.*, 2016). Fatores externos ao processo de avaliação podem também resultar em inconsistências, como por

exemplo, a capacidade de transmitir o exercício pelo entrevistador. No caso do exercício da troca de tempo, há evidências importantes sobre efeito do entrevistador na probabilidade de ocorrência de valorações inconsistentes (BADIA, MONSERRAT e HERDMAN, 1999).

O conceito de inconsistência lógica é complexo, especialmente quando o objetivo do estudo é a obtenção de preferências individuais. Existem evidências sobre indivíduos que já vivenciaram condições de saúde debilitadas e alteraram a forma de valorização de determinados estados e sua resistência a piores condições de saúde se elevou (DOLAN, 1996). Pessoas com experiência na área da saúde, que possuem algum tipo de doença crônica ou que já sofreram com algum tipo de restrição mais aguda podem valorizar estados de saúde mais debilitados de forma mais leve do que o restante da população.

Diante da relevância do fenômeno de inconsistências lógicas, recorrente na literatura, este trabalho busca estudar mais profundamente a incidência destas em dois estudos distintos aplicados à população brasileira. O desenho do estudo apresentou uma diferença importante na escolha dos estados de saúde aplicados nas duas amostras. A ideia é explorar como esta diferença metodológica afetou a consistência dos entrevistados e verificar como as respostas foram influenciadas por este desenho de estudo.

A presença de inconsistências lógicas no processo de valoração pode indicar que o respondente não compreendeu a atividade proposta ou que não percebeu a relação de dominância entre os estados apresentados. Andrade *et al.* (2015) investigam mais a fundo a presença das inconsistências lógicas numa pesquisa aplicada no estado de Minas Gerais. Os autores discutem que a comparação entre estados de saúde muito próximos pode exigir uma percepção de mudança muito sutil por parte do entrevistado.

Craig e Ramachandran (2006) apontam que os incentivos aos participantes da pesquisa devem ser explicitados para que a ocorrência de desvios do comportamento seja reduzida. Caso não fique claro o objetivo do estudo, o agente pode ter incentivos a burlar a pesquisa por diversas razões. As pesquisas de obtenção de preferências sociais por saúde costumam ser extensas, cognitivamente desafiadoras e complexas, podendo levar à confusão ou ao enfado do

respondente. O nível de educação e a familiaridade com pesquisas são fatores que podem afetar a quantidade de respostas inconsistentes e, conseqüentemente, a qualidade do estudo.

O método de aplicação da pesquisa também tem impacto sobre os resultados dos estudos. Pesquisas postais não assistidas apresentaram inconsistências lógicas com maior frequência do que as pesquisas feitas face-a-face (Dolan e Kind, 1996; Ohinma e Sintonen, 1999; Devlin *et al.*, 2000). A presença do entrevistador treinado é um atributo desejável para reduzir a incidência de inconsistências lógicas. Nas pesquisas postais, a taxa de resposta fica em torno de 50% podendo apresentar um viés de seleção amostral. Uma hipótese levantada por Devlin (2000) é que a amostra de respondentes é composta em sua maioria por indivíduos que compreenderam a pesquisa e que já eram familiares com pesquisas científicas.

O treinamento dos entrevistadores se mostrou também fator significativo sobre a frequência de respostas inconsistentes observadas. Badia, Monserrat e Herdman (1999) verificaram discrepância no número de respondentes inconsistentes e nas quantidades individuais de inconsistências cometidas em seu estudo aplicado em Barcelona. Sua pesquisa contou com dois entrevistadores, brevemente instruídos a aplicar a pesquisa, sendo que um deles possuía menos experiência com pesquisas do que o outro. Os resultados do entrevistador menos experiente apresentaram mais do que o dobro do número de inconsistências do que o outro, tanto na escala analógica visual quanto no ranking. Os autores verificaram diferença estatisticamente significativa sobre a quantidade de inconsistências cometidas, controlando pelo entrevistador responsável na execução da EVA.

O número de entrevistados que cometem inconsistências e as taxas relativas de inconsistência variam substancialmente dentro das amostras para métodos de obtenção diferentes. Recorrentemente o *Time Trade Off* apresenta maior porcentagem de entrevistados inconsistentes que a escala analógica visual (ANDRADE *et al.*, 2016; LAMERS *et al.*, 2006; Badia, Monserrat e Herdman, 1999). Isso decorre da maior complexidade do método e da pouca familiaridade com esta modalidade de experimento.

3.1 Efeitos do método de aplicação das pesquisas

A qualidade das pesquisas está intrinsicamente ligada à qualidade das informações coletadas para estudar determinado fenômeno. Pesquisas voltadas para estimação de preferências sociais se mostram complexas e extensas e a forma de obtenção dos dados requer especial atenção. Nesta seção, alguns estudos de estimação das preferências sociais por estados de saúde são comparados sob o ponto de vista da coleta de dados. Pesquisas postais mostraram certo viés de seleção da amostra, sujeitos que não compreendem adequadamente o estudo acabam não enviando suas respostas. Já as pesquisas realizadas com entrevistadores permitem a utilização de exercícios mais complexos, graças a explicação de um pesquisador treinado. Esta característica pode, no entanto, induzir algum tipo de viés no entrevistado. Os custos de treinamento e recrutamento de entrevistadores também são fatores relevantes na condução de estudos desta natureza, especialmente para grandes amostras.

O estudo do EQ-5D aplicado em Rotterdan (ESSINK-BOT *et al.*, 1993) se baseou numa pesquisa postal. Os autores adotaram uma regra de exclusão de questionários com mais de dois estados sem avaliação, assumindo que os respondentes que atribuíram valor aos outros 14 estados de saúde (pesquisa comparava 16 estados de saúde) haviam compreendido o exercício. A taxa de resposta neste caso foi de 70% dos questionários enviados. Os autores excluíram os questionários de indivíduos que valorizaram todos os estados da mesma forma, pior estado e melhor estado com mesmos valores e estados de saúde intermediários valorizados como superiores à saúde perfeita. Com essas medidas a amostra se reduziu a 46% dos questionários enviados. Justificaram a exclusão destes respondentes inconsistentes pela forma de normalização adotada no estudo, afirmando que se algum estado fosse superior à saúde perfeita, valores maiores do que 100 ocorreriam em sua escala, o que poderia levar a valores extremos que distorceriam suas ponderações.

Dolan e Kind (1996) estudaram o comportamento de duas amostras, uma baseada em pesquisa postal e outra baseada em entrevista face a face aplicadas na Inglaterra, ambos utilizando a escala analógica visual. Na primeira, obtiveram uma taxa de resposta relativamente baixa de 35%, apesar dos esforços de oferecer os incentivos corretos aos participantes. Na amostra de entrevistados, taxa de resposta ao estudo foi de 83%. Os autores relatam que em pesquisas postais

costuma ocorrer uma pré-seleção dos indivíduos que acreditam ter compreendido a pesquisa, levando a um possível viés de seleção. Estes resultados podem ser verificados nas de taxas médias de inconsistências mais baixas produzidas nesta amostra (9,3%), comparativamente àqueles obtidos na amostra que foi entrevistada (13,5%).

Ohinmaa e Sintonen (1999) realizaram aplicação do EQ-5D na Finlândia para estimar os parâmetros de valorização dos estados de saúde também utilizando uma pesquisa postal, baseada na escala analógica visual. Realizaram a exclusão dos respondentes que cometeram mais que três inconsistências lógicas. O estudo mostrou que incluir estes indivíduos em modelos para estimar os parâmetros de valorização dos estados de saúde poderia afetar sensivelmente as estimativas dos pesos sociais.

Devlin *et al.* (2000) conduziram a pesquisa do EQ-5D na Nova Zelândia, obtendo uma taxa de resposta de 45% na pesquisa postal. Verificaram que 80% dos respondentes cometeram ao menos uma inconsistência lógica. Dentro das 1360 respostas obtidas no estudo, excluíram aquelas que inviabilizariam a estimação dos parâmetros de valorização dos estados de saúde (poucos estados avaliados, mesmo valor a todos os estados, não avaliação da saúde perfeita ou da morte), ficando com 919 questionários adequados.

Uma característica em favor das pesquisas assistidas é a possibilidade de aplicação de exercícios baseados na teoria da escolha, como o TTO e o SG. A aplicação deste tipo de exercício requer um entrevistador treinado ou alguma interface que seja acessível ao usuário, como um software autoexplicativo (BANSBACK *et al.*, 2012).

3.2 Evidências empíricas internacionais

Krabbe, Essink-Bot e Bonsel (1997) conduziram um estudo para testar a confiabilidade de métodos alternativos de obtenção dos parâmetros de valorização dos estados de saúde na Holanda. Os autores compararam os exercícios da EVA, do TTO e do SG, obtendo taxas de inconsistências de 2%, 4,3% e 4,6%, respectivamente. Neste estudo, os autores verificaram que

as inconsistências lógicas costumam se concentrar em estados muito próximos entre si, o que indica que os indivíduos cometem desvios lógicos ao não diferenciar esses estados.

Badia, Mosenrat e Herdman (1999) realizaram o estudo para a obtenção dos parâmetros de valorização dos estados de saúde para residentes da Catalunha, Espanha. Os autores obtiveram proporção de respostas inconsistentes de 24,4% para o ordenamento, 25,9% para a escala analógica visual e 59,2% para o TTO. Afirmam que o exercício de obtenção das preferências afeta substancialmente estes percentuais, dada a crescente dificuldade cognitiva verificada na condução do estudo. Calcularam também as taxas médias de inconsistência cometidas nos três estudos, no TTO a taxa foi de 4%, na EVA 1,32% e no exercício de ordenamento 1,14%. A idade e o grau de escolaridade (utilizado como proxy para capacidade cognitiva) foram características relevantes para explicar a probabilidade de um indivíduo cometer inconsistência lógica no TTO e na EVA. Indivíduos idosos e menos escolarizados apresentaram maiores percentuais de inconsistências lógicas nesses dois exercícios. O ordenamento não mostrou influência dessas variáveis na quantidade de indivíduos inconsistentes ou nas taxas de inconsistências.

Lamerset *al.* (2006) encontraram que 65% dos respondentes apresentaram ao menos uma inconsistência na EVA e 89% no TTO no estudo aplicado na Holanda para a obtenção das preferências sociais pelos estados de saúde. Em uma população de adultos, a média das taxas de inconsistência foi de 1,6 na EVA e 2,7 no TTO, onde os valores máximos eram 8 e 10, respectivamente. Os autores também investigaram se as inconsistências lógicas cometidas na EVA também ocorreriam no TTO, através da correlação de Pearson. Encontram correlação relativamente baixa, de 19%, o que sugere que as inconsistências cometidas em um exercício não se estenderam para o outro. As tarifas sociais produzidas para os estados de saúde foram robustas à presença dos respondentes inconsistentes, diferentemente do que sugeriram Ohinmaa e Sintonen (1999). Lamers *et al.* (2006) defendem que as respostas inconsistentes não distorceram os pesos sociais atribuídos aos estados de saúde.

Craig e Ramachandran (2006) sugerem algumas formas de identificação de sujeitos inconsistentes, que podem ser divididos entre *non-traders*, sabotadores, *shufflers* e indivíduos com avaliações insuficientes (*missings*). Os *shufflers* e os sabotadores costumam ser os mais

difíceis de se identificar na amostra e em geral são aqueles que mais prejudicam a obtenção dos pesos necessários para a construção dos AVAQ. Indivíduos que avaliam aleatoriamente os estados de saúde apresentados (*shufflers*) podem configurar confusão no entrevistado ou cansaço na coleta das informações, de modo que seu objetivo se torna apenas encerrar a pesquisa.

Os sabotadores têm um perfil distinto, seu incentivo para burlar o estudo é maior do que o incentivo a completar a tarefa. Algumas vezes podem até começar o estudo revelando suas preferências, mas ao tentar antecipar o fim da pesquisa pelo cansaço ou pelo enfado, acabam fornecendo respostas aleatórias para encerrar o estudo. As valorações extraídas de indivíduos sabotadores podem atribuir valores mais altos a estados de saúde mais graves e valores mais baixos a estados de saúde mais próximos à saúde perfeita, sem refletir suas preferências reais pelos estados.

Indivíduos que não estão dispostos a trocar longevidade por qualidade de vida sob nenhuma circunstância no SG ou TTO são conhecidos como “*non-traders*” (CRAIG e RAMACHANDRAN, 2006). Estes entrevistados podem justificar suas preferências por crenças e algumas vezes podem atribuir à morte um valor superior à saúde perfeita, argumentando que após a morte não existem mais problemas de saúde (DEVLIN, HANSEN e SELAI, 2003). Pelo fato de nunca estar dispostos a reduzir anos de vida saudável em relação à morte, sua escolha deixa de ser sob incerteza e passa a ser determinística em relação à morte, ferindo a hipótese central da teoria da escolha.

Craig e Ramachandran (2006) compararam as pesquisas nacionais de obtenção de preferências no Reino Unido (GUDEX et al, 1994) e nos Estados Unidos (SHAW, COONS e JOHNSON, 2002). Em ambas as pesquisas, afirmam que incluir os respondentes *shufflers* e sabotadores envia os resultados das tarifas dos estados de saúde, aumentando a variância e diminuindo as dispersões das estimativas. A exclusão destes indivíduos alterou a composição das amostras e os valores estimados dos estados de saúde significativamente, tanto no TTO quanto na EVA. Verificam que algumas características sociodemográficas estão correlacionadas com a ocorrência de inconsistências lógicas, tais como idade, que apresentou correlação positiva com a quantidade de inconsistências lógicas cometidas (quanto mais velho, maior a quantidade de inconsistências), e a

escolaridade (quanto maior o grau de escolaridade verificado, menor a quantidade de desvios lógicos).

Os autores ainda afirmam que não existe *trade-off* entre quantidade e qualidade das respostas sob a presença de inconsistências e são taxativos na exclusão deste tipo de resposta. Sugerem que uma nova ponderação pode diminuir o problema de seleção amostral induzido pelas exclusões. Ao invés de criar novos testes para detecção e tratamento das inconsistências lógicas, indicam o desenvolvimento de mecanismos que controlem a ocorrência destas durante a realização do estudo.

Na pesquisa nacional implementada na Dinamarca para obtenção das preferências nacionais por estados de saúde, realizada através do EQ-5D, Wittrup-Jensen, Lauridsen e Pedersen (2008) compararam os três exercícios, ordenamento, EVA e TTO. Os autores investigaram tanto a ocorrência das inconsistências lógicas quanto a consistência individual dos entrevistados entre os exercícios respondidos. Os autores verificaram proporções de inconsistentes de 60,96% para o ordenamento, 66,59% para a EVA e 78,62% para o TTO. Já com relação às taxas de inconsistência, verificaram 4,28% para o ordenamento, 4,34% para a EVA e 7,33% para o TTO. A consistência de critério, avaliada pela correlação de Kendal, apontou 60% de correlação entre as valorizações produzidas no ordenamento versus o TTO. Indivíduos idosos, pouco educados e mais pobres apresentaram maiores taxas de inconsistências nesta pesquisa.

Wittrup-Jensen, Lauridsen e Pedersen (2008) sugerem que esforços devem ser feitos para diminuir o número de inconsistências lógicas ocorridas. Os idosos devem receber atenção especial na pesquisa, a reaplicação dos instrumentos sob nova orientação pode ser uma alternativa viável para reduzir os níveis de inconsistências lógicas desta fração da população. Enaltecem também a importância de entrevistadores experientes na aplicação da metodologia, apesar de não possuírem dados suficientes para analisar o impacto deste efeito em sua amostra.

Bansback *et al.* (2012) no estudo canadense do EQ-5D, com uma amostra de 2394 indivíduos, apresentou a exclusão de 43,7% deles. As exclusões ocorreram devido à ausência de informações, indivíduos que apresentaram o mesmo valor para todos os estados de saúde, ou que cometeram

mais do que uma inconsistência lógica. Estes autores implementaram a pesquisa através de um software, argumentando que esta modalidade pode reduzir possível viés do entrevistador, reduzir erros na formulação das questões e também diminuir os erros na coleta dos dados. As desvantagens desta aplicação se assemelham àquelas observadas nas pesquisas postais, a compreensão e realização do exercício do TTO pode ficar comprometida em alguns casos.

No estudo para a França, Chevalier e Pourvoirville (2013) apontam que, dos 452 respondentes utilizados na pesquisa, 90% da amostra cometeu ao menos uma inconsistência lógica. Os autores não excluíram respondentes baseados no número de inconsistências lógicas cometidas. A principal causa das inconsistências no TTO era a falta de compreensão do exercício ou a dificuldade sua para realização.

No estudo de Taiwan para o cálculo dos valores dos estados de saúde, Lee *et al.* (2013) também realizam a exclusão de porção expressiva da amostra, utilizando apenas cerca de 62% dela. Os parâmetros para exclusão foram aqueles tipicamente utilizados, citados no início desta seção, além de respondentes que cometeram além de 4 inconsistências lógicas. Das exclusões, a maior parte é composta por indivíduos idosos, com menor escolaridade.

3.3 Estudos aplicados no Brasil

O estudo conduzido no Brasil sobre a presença de inconsistências lógicas na pesquisa implementada no estado de Minas Gerais verificou que 53% dos entrevistados cometeu ao menos uma inconsistência no ranking, 46% na EVA e 52% no TTO (ANDRADE *et al.*, 2016). As taxas médias individuais de inconsistências foram maiores do que aquelas observadas em outras pesquisas, 8,7% no ranking, 7,6% na EVA e 10,6% no TTO. A observância destes dados traz à tona os resultados verificados em outros países, em que os indivíduos menos escolarizados apresentam maiores taxas de inconsistência. Países em desenvolvimento, com níveis de escolaridade mais baixos, precisam de pesquisas que reforcem a adoção de medidas que possam reduzir os níveis de inconsistências da população.

Ainda neste estudo, os autores verificaram quais as características individuais que mais afetaram as chances de se cometer uma inconsistência lógica. Verificaram que a indivíduos mais educados cometeram menos inconsistências no ordenamento e no TTO, mas essa relação não se verificou na EVA. Entrevistados que se enquadram nas classes de riqueza mais altas também apresentaram menores chances de incorrer em desvios lógicos. A faixa etária do indivíduo apresentou significância estatística para reduzir a chance de inconsistência para o TTO. Outro resultado interessante apontado pelo estudo é que indivíduos que se auto avaliaram como felizes tiveram menores chances de cometer inconsistências lógicas, sugerindo que o equilíbrio emocional pode contribuir para a melhor realização do estudo.

Apesar das proporções de inconsistência de cerca de 50% para as três formas de avaliação, os resultados verificados em Minas Gerais não são tão discrepantes daqueles observados nos países desenvolvidos. A avaliação dos métodos de valorização dos estados de saúde se mostrou factível apesar da grande heterogeneidade socioeconômica verificada na pesquisa.

Cada indivíduo avaliou apenas 8 estados de saúde nesta pesquisa, diferentemente dos estudos europeus em que 13 estados de saúde definidos pelo EQ-5D-3L eram avaliados. Este protocolo tem dois efeitos sobre as taxas individuais de inconsistência, por um lado o aprendizado que pode ocorrer durante a elaboração do estudo é menor, devido a menor exposição a diferentes estados de saúde. Por outro lado, como a pesquisa é mais breve, as chances de fadiga ou enfado por parte do entrevistado se reduzem substancialmente.

A remoção das respostas de entrevistados inconsistentes alterou significativamente as preferências sociais por saúde estimadas com base no exercício do TTO. Parcela significativa da amostra pode acabar sendo excluída do estudo devido a este tipo de censura. Os autores sugerem que métodos mais refinados devem ser implementados nos diferentes contextos econômicos e demográficos. Novas tentativas de melhorar a consistência das respostas de indivíduos menos educados ou que vivenciam condições de vida desfavoráveis devem ser desenvolvidas para captar as preferências desta fração social. Enquanto estes métodos não são implementados, métricas mais precisas para a classificação dos indivíduos inconsistentes também podem ser desenvolvidas.

Neste trabalho, utilizar-se-á os dados de duas pesquisas realizadas no Brasil que permitiu estimar os parâmetros de valoração para os 243 estados de saúde do EQ-5D-3L. A primeira pesquisaimplementada em Minas Gerais foi realizada entre outubro e dezembro de 2011. A segunda pesquisa foi realizada nas cidades de Porto Alegre, Rio de Janeiro e Recife em 2012. Ambas seguiram protocolos similares mas apresentaram duas importantes diferenças. A pesquisa implementada no estado de Minas Gerais utiliza um desenho de estudo baseado em blocos fixos de estados de saúde, seguindo uma revisão do protocolo MVH proposta por Kind (2009) onde os 102 estados de saúde diretamente avaliados são classificados em 26 blocos com cada um contendo 2 estados leves, 2 moderados e 2 graves (blocos balanceados). Na pesquisa realizada nas três cidades foram avaliados diretamente os 243 estados de saúde e cada indivíduo avaliou um bloco de estados de saúde distribuídos de forma aleatória. Dessa forma, esse estudo não garantiu que cada indivíduo avaliasse pelo menos um estado de saúde leve, um moderado e um grave. O objetivo dessa dissertação é comparar a presença de inconsistências entre esses dois estudos. A hipótese a ser testada é se a valoração de blocos de estados de saúde não balanceados poderia acarretar em maior nível de inconsistência.

A ocorrência de inconsistências lógicas deve sofrer impactos relevantes decorrentes dessa diferença metodológica. Como o balanceamento proposto dentro da estrutura de blocos fixos delimita uma relação mais clara entre a gravidade dos estados de saúde, espera-se obter taxas de inconsistências mais baixas do que na pesquisa que não utiliza os blocos fixos (não balanceados).

Outra contribuição proposta neste trabalho visa expandir a análise das inconsistências lógicas levando em conta o conceito da distância entre estados de saúde tal como proposto em Andrade et al (2015). O objetivo é testar se a ocorrência dos desvios lógicos se dá entre estados muito próximos, indicando que os respondentes não foram capazes de perceber a diferença entre estados com diferenças marginais, ou se as inconsistências ocorrem entre estados muito distantes. Novamente o desenho do estudo é um fator relevante, uma vez que a estrutura de blocos fixos leva em conta a distância entre estados de saúde na sua construção. Espera-se que o desenho do estudo balanceado e não balanceado tenha impactos sobre a quantidade de inconsistências

individuais cometidas e também sobre a distribuição destas inconsistências conforme as distâncias entre os estados de saúde variam.

4. Aspectos Metodológicos

4.1 Bases de dados

Este trabalho compara as duas pesquisas implementadas no Brasil para obtenção dos parâmetros de valorização dos estados de saúde. O primeiro estudo, projeto “Estimação dos Parâmetros de Valorização dos Estados de Saúde em Minas Gerais a partir do EQ-5D”, foi realizado pelo CEDEPLAR/UFMG (Centro de Desenvolvimento e Planejamento Regional) com financiamento da FAPEMIG (Fundação de Amparo à Pesquisa do estado de Minas Gerais), no âmbito do edital Programa Pesquisa para o SUS – PPSUS. A pesquisa foi conduzida em 2011 e coordenada pelas professoras Mônica Viegas e Kenya Noronha. O segundo projeto expandiu a pesquisa de Minas Gerais para outras três cidades com o objetivo de obter uma amostra que incluísse outros estados do Brasil para estimação de preferências sociais que fossem utilizadas pelo Ministério da Saúde como parâmetros nacionais. O projeto teve a coordenação geral do NATS/INC– Núcleo de Avaliação de Tecnologias em Saúde/ Instituto Nacional de Cardiologia, em parceria com o Centro de Desenvolvimento e Planejamento Regional – CEDEPLAR/UFMG, a Universidade Federal do Rio Grande do Sul - UFRGS e a Universidade Federal de Pernambuco UFPE. Na estimação dos parâmetros nacionais, foram incluídas, além do estado de Minas Gerais, as cidades de Porto Alegre, Recife e Rio de Janeiro (SANTOS *et al.*, 2015).

Embora as duas pesquisas tenham sido realizadas com o objetivo de obter as preferências sociais por estados de saúde para o Brasil, algumas mudanças no desenho do estudo foram introduzidas na pesquisa das três cidades. Os dois estudos adotaram o mesmo protocolo de entrevista, seguindo uma adaptação do protocolo original do MVH *Measurement and Valuation of Health*, (GUDEX, 1994), sendo, portanto, factível a sua comparação. O trabalho aqui proposto aproveita essa oportunidade inédita, de comparar os resultados obtidos em dois estudos que seguem o mesmo protocolo, explorando algumas mudanças metodológicas introduzidas no estudo das três cidades. A principal diferença metodológica no desenho de estudo entre as duas pesquisas, concerne a definição dos estados avaliados. No estudo de Minas Gerais, 102 estados de saúde

foram definidos em blocos fixos de 09 estados balanceados de acordo com a gravidade (02 estados grave, 02 estados moderados e 02 estados leves, saúde perfeita e o pior estado de saúde possível) e avaliados pela população entrevistada. No segundo estudo, todos os 243 estados de saúde definidos pelo EQ-5D-3L foram avaliados. A distribuição dos estados na população não seguiu uma estrutura de blocos fixos e foi randomicamente distribuída de modo apenas a obter um número mínimo de avaliações por estados de saúde. Com esse desenho, o grupo de estados avaliados por cada indivíduo no estudo Nacional não foi necessariamente balanceado como na maior parte dos estudos (GUDEX *et al.*, 1994; SHAW, COONS e JOHNSON; 2002; GREINER *et al.*, 2005; WITTRUP-JENSEN; LAURIDSEN e PEDERSEN; 2008; CHEVALIER e POURVOURVILLE; 2013). Diversos autores defendem o uso da estrutura de blocos de estados de saúde balanceado principalmente para facilitar a comparabilidade entre os estados de saúde, evitando a presença de inconsistências e aumentando a eficiência na estimação das preferências por estados de saúde (GREINER *et al.*, 2005; WITTRUP-JENSEN; LAURIDSEN e PEDERSEN; 2008; KIND; 2009)

4.1.1 Amostra e coleta de dados do estudo de Minas Gerais

O estado de Minas Gerais apresenta grandes disparidades em termos de qualidade de vida e de desenvolvimento econômico. O índice de desenvolvimento humano municipal (IDH-M) de 2000 apresentou resultados que variaram de 0,57 na região nordeste do estado até 0,84 na região sudeste. A heterogeneidade presente no estado de Minas Gerais, em alguma medida, é similar àquela observada no Brasil, onde o intervalo de variação do IDH estadual em 2000 foi de 0,64 (Maranhão) a 0,82 (Santa Catarina) (PNUD; 2003). A grande diversidade de indicadores socioeconômicos presente no estado pode ser representativa da heterogeneidade socioeconômica da população brasileira, viabilizando, com as devidas proporções, extrapolações dos resultados obtidos no estado para a população nacional.

A amostra do estado de Minas Gerais incluiu informações de indivíduos alfabetizados, com idades entre 18 e 64 anos, que moravam em regiões urbanas do estado. A definição do tamanho amostral foi baseada nas informações do Censo Demográfico de 2010 com margem de erro de

3%. Foram entrevistados 3362 indivíduos, distribuídos entre Belo Horizonte (1115), região metropolitana da capital (626) e interior do estado (1621) em áreas urbanas. A amostra é probabilística no nível de setor censitário para a cidade de Belo Horizonte e estratificada por sexo e idade. Para a Região Metropolitana e interior do estado a amostra é probabilística no nível de município. A pesquisa é representativa da população urbana do estado de Minas Gerais, para a Região Metropolitana de Belo Horizonte e para a capital do estado.

A distribuição das características demográficas e socioeconômicas da amostra evidenciam perfil similar ao observado em outras pesquisas domiciliares conduzidas para a população do estado de Minas Gerais, como a PAD (Pesquisa por Amostras de Domicílios) e a PNAD (Pesquisa Nacional por Amostra Domiciliar). A tabela 5 apresenta a comparação para algumas características populacionais obtidas nestas pesquisas. A distribuição das variáveis de idade e sexo é bem próxima aos demais estudos, o que valida a amostra como representativa do estado.

O intervalo de idade de 18 a 64 anos foi definido para obter preferências considerando o maior grupo populacional possível, sendo a idade média dos entrevistados de 38 anos. A amostra exclui os idosos e as crianças devido a maior complexidade na aplicação dos exercícios para esses grupos. Para as crianças, recomenda-se o uso do sistema descritivo específico definido como EQ-5D-Y (WILLE *et al.*, 2010). A maior parte dos estudos que objetivam a estimação das preferências sociais por estados de saúde incluem na amostra os indivíduos acima de 18 anos de idade (OHINMAA e SINTONEN, 1999; GUDEX *et al.*, 1994; BADIA, HERDMANN e MONTSERRAT, 1999; DEVLIN *et al.*, 2000).

A exclusão de idosos acima de 64 anos no caso brasileiro decorre do elevado nível de analfabetismo observado nesse grupo etário no Brasil. Segundo os dados da PAD, em 2013, 26,5% dos idosos acima 60 anos eram analfabetos enquanto entre os adolescentes de 15 a 19 anos essa taxa era de apenas 0,9%. Essa diferença resulta do aumento ao acesso ao ensino no Brasil nas últimas décadas (FJP, 2015). O protocolo de valoração dos estados de saúde por exigir capacidade cognitiva dos entrevistados é inviável de ser aplicado para indivíduos analfabetos.

Em relação a presença de plano de saúde, o percentual observado na amostra do estudo é similar ao observado para as outras pesquisas disponíveis para Minas Gerais: cerca de 30% da população do estado dispõe do serviço privado (31,36% na PPSUS, 28,34% na PAD e 35,38% na PNAD). A variável de estado saúde auto reportado também apresentou distribuição similar nas três pesquisas com predomínio do estado de saúde muito bom ou bom apenas cerca de 3% dos indivíduos respondem ter saúde ruim ou muito ruim nas pesquisas.

Foram recrutados 25 entrevistadores, 2 auxiliares de coordenação e cinco especialistas em revisão e codificação. O treinamento dos pesquisadores foi feito pela equipe do GEESC/UFMG durante três dias. Foram aplicados questionários-teste aos entrevistadores para garantir maior eficiência em campo. Em decorrência da grande complexidade da pesquisa, a equipe de entrevistadores se reduziu a 20, com desistência de 5, e ainda 7 foram reprovados mesmo após terem sido novamente treinados. Foram confeccionados coletes aos pesquisadores e 28 Quadros de Troca de Tempo, conforme previsto no protocolo adotado.

Tabela 5. Características sociodemográficas da amostra de Minas Gerais e comparação outras bases de dados

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).

As entrevistas domiciliares foram realizadas no período de outubro a dezembro de 2011, com uma média diária de 5 entrevistas diárias por pesquisador. Em cada domicílio foi entrevistado somente um morador para evitar problemas de correlação serial entre os indivíduos. Não foram oferecidos incentivos econômicos aos respondentes e o tempo médio de entrevista foi de 45 minutos. Os primeiros lotes da pesquisa apresentaram alguns equívocos que serviram de base para uma nova rodada de treinamentos. A equipe responsável pela coordenação realizou revisões e checagens para confirmar a autenticidade dos resultados, cerca de 20% dos questionários foram

revisados por telefone para atestar a qualidade da pesquisa e minimizar a chance de fraudes pelos entrevistadores.

4.1.2 Amostra do estudo das três cidades

O segundo estudo investigou indivíduos residentes das cidades de Porto Alegre, Rio de Janeiro e Recife. Esta amostra contempla 5785 indivíduos sendo 3932 na cidade do Rio de Janeiro, 894 em Porto Alegre e 959 em Recife. As características demográficas e socioeconômicas da amostra encontram-se na tabela 6 que compara a amostra das três cidades com a de Minas Gerais, O desenho amostral da pesquisa realizada nas cidades de Porto Alegre, Recife e Rio de Janeiro foi baseada nos dados do Censo 2010. As informações foram coletadas para uma amostra probabilística por setor censitário da população urbana, incluindo indivíduos de 18 a 64 anos e, estratificada por idade e sexo. A amostra é representativa da população urbana adulta das cidades do Rio de Janeiro, Porto Alegre e Recife. Observa-se proporção superior de homens (53,77%) comparada à amostra de MG. A distribuição etária apresenta maior concentração na faixa entre 50 e 59 anos (18,34% ante 16,25% no estado). Em relação a escolaridade, nas três cidades, a distribuição é mais concentrada nos graus primário e fundamental completos (30,70% e 37,52%, respectivamente) e havendo menor participação no ensino médio completo (22,55% ante 37,64% no estado de Minas Gerais).

A posse de plano de saúde é bastante diferente nas duas amostras: na amostra das três cidades, 44% dos respondentes afirmaram ter plano de saúde, número superior ao observado para Minas Gerais, 31,36%. Em relação a saúde auto avaliada, a amostra de Minas Gerais reporta ser relativamente mais saudável do que a das três cidades: 77% dos indivíduos mineiros reportam ter saúde boa ou muito boa, enquanto que nas três cidades esse percentual é de 68%.

A pesquisa de campo foi realizada ao longo do ano de 2012 em períodos intercalados para compatibilizar as atividades de treinamento e supervisão dos entrevistadores. No Rio de Janeiro a coleta de dados ocorreu entre 28 de abril e 30 de outubro; em Porto Alegre entre 19 de maio e 24 de setembro e no Recife de 19 de junho a 1º de outubro. Foram aplicados 4227 questionários no

Rio de Janeiro, 964 em Porto Alegre e 1032 em Recife. No Rio de Janeiro foram empregados 33 entrevistadores, 10 em Porto Alegre e 15 em Recife. Nas três capitais foram realizados três dias de treinamento para os entrevistadores, visando garantir a qualidade dos dados coletados nesta pesquisa.

Para verificar a qualidade das entrevistas, foram selecionados aleatoriamente alguns questionários e feita auditoria por telefone para checar a veracidade das informações obtidas. Foram realizadas exclusões prévias de questionários em que o exercício de troca de tempo foi aplicado incorretamente, exclusões devidas a cotas de sexo e idade e por fraude.

Alguns desafios que a equipe responsável reportou foram a receptividade da população a esta nova forma de coleta de preferências por estados de saúde, distinta de outros exercícios mais simples. A compreensão dos questionários era uma questão primordial para a qualidade dos dados, para que isso fosse alcançado foi necessário o re-treinamento das equipes de entrevistadores, o que também elevou a taxa de rotatividade nas equipes. Outro fator relevante encontrado foi o analfabetismo funcional entre os respondentes, que prejudica a compreensão e a diferenciação dos estados de saúde pelos entrevistados, comprometendo a qualidade dos dados da pesquisa e gerando um grande número de inconsistências lógicas.

Tabela 6. Características Socioeconômicas Amostra Nacional

Características	Minas Gerais	Amostra Nacional
Sexo (%)		
Homem	51,58	53,77
Mulher	48,42	46,23
Faixa Etária (%)		
18-34	43,3	42,75
35-49	33,95	32,53
50-59	16,25	18,38
>60	6,5	6,33
Nível Educacional (%)		
Prim. Inc.	4,86	2,79
Prim. Comp.	24,37	30,7
Fund. Comp.	24,56	37,52
E.M. Comp.	37,64	22,55
Sup. Comp.	8,54	6,45
Estado Civil (%)		
Casado	56,12	50,67
Viúvo	2,83	3,16
Divorciado	7,92	8,43
Solteiro	33,13	37,69
Plano de Saúde (%)		
Sim	31,36	44,24
Não	68,64	55,67
Saúde Auto-Avaliada (%)		
Muito Bom	25,35	18,95
Bom	52,01	49,62
Regular	20,49	28,37
Ruim	1,58	2,37
Muito Ruim	0,49	0,66

FONTE: Elaboração do autor

Prim. Inc.: ensino primário incompleto; Prim.Comp: ensino primário completo; Fund.Comp.: ensino fundamental completo; E.M. Comp.: ensino médio completo; Sup. Comp.: ensino superior completo

*Amostra nacional excluindo a fração mineira.

4.2 Aplicação do protocolo da entrevista

A aplicação dos questionários da pesquisa, para ambos os estudos, utilizou três instrumentos, denominados A, B e C, apresentados no anexo A deste trabalho. Estes instrumentos seguem o protocolo do MVH com algumas adaptações. No instrumento A, realizado no início da entrevista cada indivíduo reporta seu estado de saúde atual, a partir do EQ-5D-3L e atribui um valor a seu estado de saúde a partir da escala analógica visual, no intervalo de 0 – 100, sendo 0 o pior estado de saúde possível e 100 a saúde perfeita. Em seguida os indivíduos realizam o exercício do

ordenamento. Nesse caso, o indivíduo recebe 9 cartões contendo estados de saúde descritos pelo EQ-5D-3L e deve ordená-los de acordo com suas preferências. Os nove cartões, contem 06 estados de saúde, o pior estado (33333), a morte e o estado de saúde perfeita (11111). A seguir, o entrevistador deve recolher os cartões e devolver ao respondente de forma aleatória, para a realização do exercício da escala visual analógica. No protocolo do MVH os indivíduos realizavam os dois exercícios sem que os cartões fossem embaralhados pelo entrevistador. O entrevistado deve atribuir um valor para cada estado de saúde na escala analógica, considerando a informação que deve se imaginar nesta condição de saúde por 10 anos seguida pela sua morte. Depois da realização destes exercícios, os indivíduos já estão familiarizados com a descrição dos perfis de saúde, e estão aptos para a execução do exercício de troca de tempo.

No instrumento B, o entrevistado deverá realizar escolhas que envolvem longevidade e qualidade de vida, através do exercício do *Time Trade Off*. O indivíduo deverá valorizar sete cartões neste exercício, os seis 6 estados de saúde já avaliados nos exercícios anteriores do ordenamento e escala visual analógica mais o pior estado de saúde possível (33333). As âncoras do exercício da troca de tempo são a saúde perfeita e a morte, os cartões que contém estes dois estados de saúde são recolhidos pelo entrevistador. A realização do exercício é conduzida com o auxílio de um quadro dupla-face, também apresentado no anexo B deste trabalho. Cada perfil de saúde é valorizado separadamente. Inicialmente o indivíduo avalia se o estado de saúde apresentado é melhor ou pior do que a morte imediata. Se o estado for melhor do que a morte, o indivíduo procede ao exercício usando o lado A do quadro do exercício da troca de tempo. No quadro A o indivíduo realiza uma troca entre número x de anos ($x < 10$) sob a saúde perfeita que equivalem a viver com a condição descrita no estado de saúde por 10 anos. Em outras palavras, o indivíduo troca tempo por qualidade de vida. Quanto menor o valor de x , pior é considerado o estado de saúde pelo indivíduo uma vez que está disposto a trocar mais tempo por qualidade. No caso de o indivíduo avaliar o estado de saúde como pior que a morte o exercício da troca de tempo é realizado utilizando a face B do quadro. Nesse caso, os indivíduos escolhem morrer imediatamente ou viver por um número x de anos ($x < 10$) na condição de saúde descrita pelo estado avaliado, 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 informações referentes às características demográficas e socioeconômicas de cada respondente

4.2.1- Diferenças no Desenho dos Estudos

Embora os dois *surveys* implementados no Brasil tenham usado o mesmo protocolo de entrevista, algumas mudanças metodológicas foram incorporadas na definição do desenho do estudo das três cidades. O trabalho aqui proposto, explora essa oportunidade de comparação de dois estudos realizados no mesmo contexto socioeconômico e que utilizam o mesmo protocolo de entrevista, mas que apresentam desenhos de estudo diferenciados. O desenho do estudo de Minas Gerais seguiu as sugestões indicadas por Kind (2009), em especial no que tange à escolha dos estados de saúde utilizados. Diferentemente do estudo original MVH, os indivíduos valoraram 102 estados de saúde, subdivididos em 26 blocos fixos, com 6 estados em cada. Cada indivíduo analisou um grupo de 9 estados, composto pela saúde perfeita (11111), “morte imediata”, o pior estado de saúde (33333), 2 estados leves, 2 estados moderados e 2 estados graves. Essa categorização dos estados de saúde segue a estrutura de blocos fixos, sugerida por Kind (2009). O grau de gravidade do estado de saúde é definido a partir da distância euclidiana entre os estados de saúde e a saúde perfeita.

A distância é uma métrica algébrica de subtração de cada dimensão do estado de saúde em relação à saúde perfeita. A distância entre o estado 22222 e a saúde perfeita é de 5, enquanto que a distância entre o pior estado de saúde (33333) e a saúde perfeita é máxima, de 10 unidades. Um exemplo de distância entre um par de estados de saúde pode ser visto entre os estados “11233” e o “11122”, toma-se a distância euclidiana, dimensão a dimensão: $(1-1) + (1-1) + (2-1) + (3-2) + (3-2) = 3$. A distribuição de todos os estados de saúde oriundos do EQ-5D-3L em termos de distância à saúde perfeita segue a tabela 7.

A distribuição da distância entre os estados e a saúde perfeita cresce monotonicamente até a distância 5, que concentra 21,5% dos estados. A partir daí as distâncias decrescem monotonicamente até a distância 10, máxima, entre o par de estados 33333 e 11111. Quanto

maior a distância, pior é o estado de saúde. Kind (2009) categoriza os estados de saúde a partir da distância classificando-os em moderados, leves e graves que segue no quadro 1.

Tabela 7. Distribuição das distâncias entre estados de saúde e a saúde perfeita

Distância	Freq Total	%	Freq MG	%
0	1	0.40%	1	0.98%
1	5	2.10%	5	4.90%
2	15	6.20%	10	9.80%
3	30	12.30%	15	14.71%
4	45	18.50%	12	11.76%
5	51	21.00%	16	15.69%
6	45	18.50%	12	11.76%
7	30	12.30%	14	13.73%
8	15	6.20%	11	10.78%
9	5	2.10%	5	4.90%
10	1	0.40%	1	0.98%
Total	243	100.00%	102	100.00%

Fonte: Adaptado de Kind (2009)

Quadro 1 -Descrição do Grau de Comprometimento dos estados de saúde

Severidade	Descrição
Estados Brandos	aqueles pertencentes aos grupos de distância entre 1 e 4, em que não existe nenhuma dimensão com problemas graves (nível 3). Esta categoria comporta 25 estados de saúde (4 dimensões nível 1 e 1 nível 2; 3 dimensões nível 1 e 2 nível 2; 2 dimensões nível 1 e 3 nível 2).
Estados Moderados	estados cujas distâncias não comportem nenhuma dimensão no nível 1 e pelo menos 2 dimensões no nível 3. Inclui um total de 25 estados (4 dimensões no nível 3 e uma no nível 2, 3 dimensões no nível 3 e 2 no nível 2, 2 dimensões no nível 3 e 3 dimensões no nível 2).
Estados Severos	compõe os demais estados, que não são severos nem brandos, formando um grupo de 50 estados.

Fonte: Adaptado de Kind (2009)

Os 26 grupos foram compostos por blocos de seis estados de saúde além do pior estado de saúde (33333), a morte e saúde perfeita. Os seis estados de saúde foram selecionados de modo a obter 2

estados graves, 2 estados moderados e 2 leves. Para a definição dos 102 estados de saúde, Kind (2009) exclui os denominados estados de saúde implausíveis. Essa estrutura de escolha dos estados de saúde, determina para todos os indivíduos, blocos de estados balanceados, permitindo, que no exercício de avaliação os indivíduos experimentem os três graus de gravidade e todas as possíveis interações entre grau de gravidade das dimensões de saúde. O balanceamento permite que os indivíduos tenham maior capacidade de distinguir os estados de saúde (SHAW, COONS e JOHNSON, 2002; WITTRUP-JENSEN, LAURIDSEN e PEDERSEN, 2008). Esse desenho de estudo permitiu a avaliação de 102 estados, ao invés de apenas 42 estados de saúde realizada no protocolo original MVH e em diversos trabalhos. O maior número de estados avaliados permite que a estimação dos valores dos demais estados seja realizada com maior eficiência através das técnicas econométricas

Tabela 8. Estados de saúde utilizados na pesquisado Estado de Minas Gerais

Estados âncora	11111; 33333
Estados brandos	11112; 11121; 11211; 12111; 21111 11122; 11131; 11113; 21133; 21222; 21312; 12211; 11133; 22121; 12121; 22112; 11312
Estados Moderados	13212; 32331; 13311; 22122; 12222; 21323; 32211; 12223; 22331; 21232; 32313; 22222
Estados severos	33232; 23232; 23321; 13332; 22233; 22323; 32223; 32232; 33321; 33323; 23313; 33212

Fonte: Adaptado de Kind (2009)

Cada estado foi avaliado por no mínimo 100 indivíduos, seguindo as recomendações de estudos anteriores (CHUANG E KIND, 2010). No exercício da Troca de Tempo cada indivíduo avaliou os 6 de estados de saúde além do 33333 (pior estado de saúde), sendo âncoras os cartões 11111 (saúde perfeita), e o cartão 'Morte', configurando ao todo 9 cartões por indivíduo.

O desenho do estudo das três cidades se diferencia da pesquisa em Minas Gerais em dois elementos chaves: 1) Definição dos estados a serem avaliados; 2) definição do grupo de estados avaliados por cada indivíduo. Neste estudo, um dos objetivos metodológicos, foi testar a

eficiência da estimação de parâmetros de valorização dos estados de saúde utilizando um desenho que inclui todos os 243 estados de saúde. A escolha por um subconjunto de estados de saúde tem vantagens e desvantagens. Entre as principais vantagens vale a pena mencionar pelos menos duas. A primeira se refere a dificuldade de conduzir estudos com tamanho amostral para obter número de avaliações suficiente de cada estado de saúde. A segunda vantagem se refere a presença de estados de saúde implausíveis. A combinação das cinco dimensões associada aos três níveis de gravidade pode resultar em alguns casos, em estados de saúde, denominados implausíveis. Esses estados se caracterizam por situações em que há relativa incompatibilidade entre as gravidades associadas as dimensões. Para os indivíduos, avaliar estados de saúde implausíveis é uma dificuldade concreta, que pode inclusive ampliar a presença de inconsistências na valoração dos estados de saúde (BAGUST, 2013). Do ponto de vista estatístico, a valoração estimada e obtida com os efeitos marginais, minora o problema de estados implausíveis. A seleção de um subconjunto de estados de saúde, por outro lado, tem como principal desvantagem a necessidade de estimação estatística dos parâmetros.

A segunda diferença importante no desenho de estudo das três cidades se refere ao grupo de estados selecionado para cada indivíduo. No caso do estudo de Minas Gerais, como já explicado anteriormente, foram definidos 26 blocos fixos com 06 estados cada. Esses 26 blocos foram randomicamente distribuídos entre os entrevistadores que deveriam cumprir cotas de idade e sexo em cada bairro/município selecionado. No estudo Nacional, os grupos de estados de saúde foram aleatoriamente gerados de forma a garantir um número mínimo de 100 avaliações de cada estado de saúde. A escolha aleatória dos estados de saúde avaliados por indivíduos, gera, portanto, um desenho que denominado doravante neste trabalho de desbalanceado da escolha dos cartões. O desenho é desbalanceado por não apresentar uma estrutura de blocos fixos e por não apresentar para cada indivíduo todos os graus de gravidade e a possível interação entre dimensões.

Cada estado de saúde recebeu ao menos 139 avaliações, e no máximo 5756 avaliações (cartão 33333). Este foi um dos primeiros estudos que realiza a estimação das preferências sociais por estados de saúde com a valoração de todos os estados de saúde oriundos do EQ-5D-3L usando o exercício da Troca de Tempo, o estudo pioneiro utilizou a escala analógica visual numa pesquisa postal no Reino Unido, conduzido por Macran e Kind (2000) e ficou conhecido como estudo de

saturação ao investigar valores para os 243 estados de saúde. Um objetivo deste trabalho é verificar se este protocolo diferenciado induz diferenças significativas na consistência lógica dos indivíduos.

4.3 Comparabilidade e Inconsistências Lógicas

Os estados de saúde são considerados comparáveis quando existe uma relação lógica de dominância entre eles. Como os estados de saúde descritos pelo EQ-5D-3L compõem um sistema de classificação nominal, pares de estados de saúde podem ser ordenados de acordo com suas dimensões de saúde. Por exemplo, o estado 33111 dominará o estado 33112 já que as dimensões apresentam graus de gravidade inferiores ou iguais àqueles observados no primeiro estado.

A fim de formalizar este conceito, faz-se necessária uma abordagem mais técnica. Como cada estado de saúde contém 5 algarismos, cada um deles representando o nível de gravidade da respectiva dimensão, o código deverá ser assim definido:

$L_{mob}L_{cdp}L_{ath}L_{ded}L_{ans}$, onde os subscritos *mob*, *cdp*, *ath*, *ded* e *ans* representam as dimensões mobilidade, cuidados pessoais, dor ou desconforto e ansiedade ou depressão, respectivamente. Cada um dos algarismos poderá assumir valores 1, 2 ou 3.

Tomando os estados de saúde hipotéticos *i* e *j*, como existem 3 graus de gravidade, denominados *k*, um par de estados será definido comparável ($C_{ij}=1$) se:

$$C_{ij} = \begin{cases} 1, & \text{se } L_k^i - L_k^j \geq 0 \forall k \cup L_k^i - L_k^j \leq 0 \forall k \\ 0 & \text{se } L_k^i - L_k^j \geq 0 \cap \exists \text{ ao menos um } k \mid L_k^i - L_k^j < 0 \end{cases}$$

Pela relação de dominância entre os estados definida por este pré-ordenamento lógico, temos:

$$i < j \text{ se } L_k^i - L_k^j \geq 0$$

$$j < i \text{ se } L_k^i - L_k^j \leq 0$$

Isto é, se os níveis de gravidade do estado i forem superiores aos níveis de gravidade do estado j (todas as dimensões de saúde do estado i apresentam tantos ou mais problemas que as dimensões do estado j), então o estado j deverá ser preferível a i . seja V_i^n o valor que o indivíduo n atribui ao estado de saúde i , define-se uma inconsistência lógica IL no par de estados de saúde (i, j) quando:

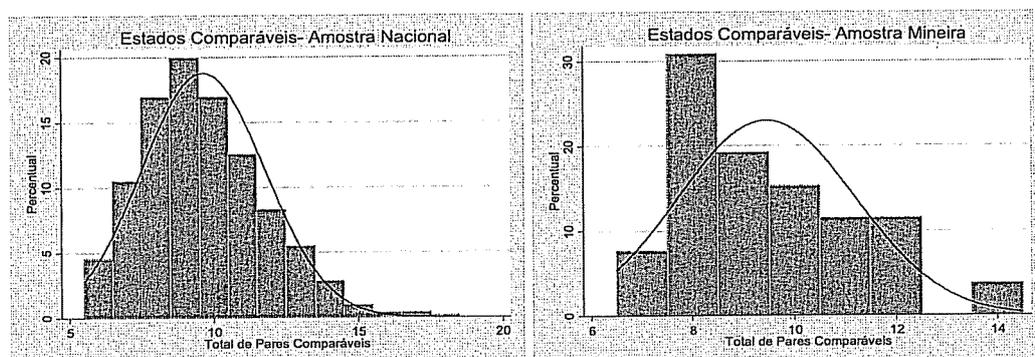
$$IL_{i,j}^n = \begin{cases} 1 & \text{se } [(V_i^n > V_j^n) \cap (L_k^i > L_k^j)] \cup [(V_i^n < V_j^n) \cap (L_k^i < L_k^j)], \\ 0 & \text{caso contrário} \end{cases} \quad \forall k$$

Assim, 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 .

A comparabilidade estará violada quando os níveis de gravidade dos estados não possuírem uma relação de dominância para todas as dimensões de saúde. Por exemplo, os estados (A) 11232 e (B) 11323 não são comparáveis pois a terceira dimensão do estado A é menos severa que a de B, enquanto a quarta dimensão de A é menos severa que a de B. Dessa forma, como não há uma clara relação de dominância entre todas as dimensões destes estados, portanto não existe comparabilidade entre eles.

No conjunto de todos de estados de saúde possibilitados pelo EQ-5D-3L é possível a formação de $C_{243,2} = 29.403$ pares, sendo que apenas 7533 (25,6%) possuem relação de dominância lógica entre si (CRAIG e RAMACHANDRAN, 2006). A adaptação do protocolo implementada na pesquisa realizada em Minas Gerais (KIND, 2009), reduz ainda mais o espectro de estados comparáveis, limitando-os a 546 pares possíveis na amostra de Minas Gerais (ANDRADE *et al.*, 2016). Ainda neste estudo, após a formação dos blocos de estados de saúde, como alguns pares comparáveis de estados estão presentes em mais do que um bloco, 220 pares de estados de saúde comparáveis foram avaliados pelos respondentes.

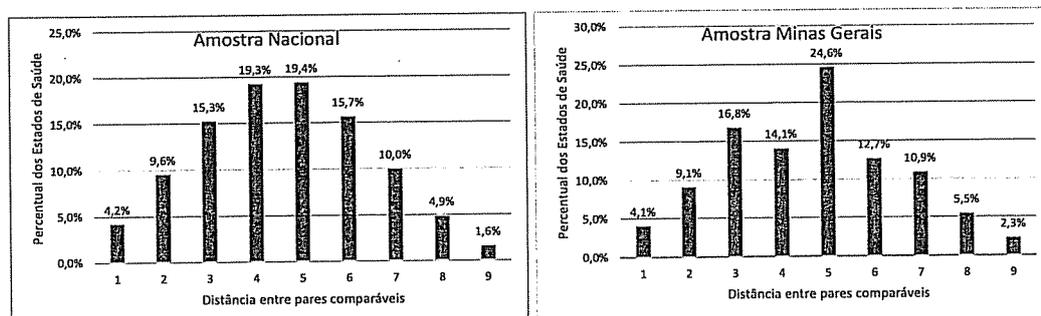
O diferente desenho de estudo proposto entre as pesquisas produz um primeiro resultado metodológico importante, a distribuição de pares de estados de saúde comparáveis entre as amostras tem comportamento distinto. A figura 1 apresenta um histograma comparativo do número de pares comparáveis disponibilizado para cada avaliação individual entre as duas amostras. A distribuição observada já reflete um resultado que afetará as demais comparações entre as pesquisas, na amostra do estado de minas gerais a distribuição se concentrou entre 7 e 14 pares de estados de saúde comparáveis, enquanto que a amostra nacional teve uma maior amplitude, de 5 a 18 pares comparáveis.



FONTE: elaboração do autor

FIGURA 3. Histograma do total de pares de estados de saúde comparáveis por amostra

A figura 2 apresenta a distribuição destes pares comparáveis levando em conta a distância entre os estados de saúde, nota-se novamente o efeito do desenho de estudo distinto aplicado entre as amostras. A amostra nacional apresenta uma distribuição mais gradual entre as distâncias, crescendo até a distância igual a 5 (acumulando 19,4% das distâncias) e então caindo gradativamente até a distância 9 (1,6% dos pares comparáveis). A amostra do estado de Minas Gerais apresenta uma distribuição com um pico na distância dos pares igual a 5 (acumulando 24,6% dos pares comparáveis), seguida por uma queda gradativa até a distância 9 (que soma 2,3% dos pares comparáveis).



FONTE: Elaboração do Autor e adaptado de ANDRADE, e al., (2015)

FIGURA4. Distribuição dos pares de estados de saúde comparáveis por amostra

A diferença na distribuição das distâncias entre os pares de estados de saúde avaliados pelos entrevistados levou também a uma diferente composição dos conjuntos de estados de saúde. Kind (2009) sugere que cada entrevistado avalie blocos que contenham 2 estados leves, dois moderados e dois graves, e esta categorização foi implementada na amostra do estado de Minas Gerais. Para fins comparativos, baseando-se na mesma categorização dos estados explicada na seção anterior, os indivíduos da amostra nacional da pesquisa ficaram distribuídos conforme o quadro 1. Observa-se que 28% da amostra nacional avaliou um bloco de 6 estados que apresentava pelo menos um estado grave, um leve e um moderado, de modo mais próximo ao que foi verificado na amostra de Minas Gerais. Cerca de 20,7% dos entrevistados realizaram valorações apenas entre estados graves e moderados, 23,91% receberam estados moderados e leves e 27,43% recebeu apenas cartões de estados de saúde moderados.

Quadro 2: Distribuição dos entrevistados pelas categorias de estados de saúde avaliados

Categorias de Estados avaliados	%
Graves e Moderados e Brandos	28,01
Graves e Moderados	20,66
Moderados	27,43
Moderados e Brandos	23,91

FONTE: Elaboração do autor

A sub-amostra que recebeu apenas estados de saúde moderados pode ter problemas em diferenciar e atribuir valores nesta categoria. Segundo Kind (2009) esta categoria é composta por estados de saúde que não são leves nem graves, totalizando cerca de 191 estados de saúde. Neste

amplo conjunto, os indivíduos podem ter mais dificuldades em diferenciar os estados, graças à heterogeneidade deste conjunto, além de conter estados de saúde que podem ser pouco factíveis (ou implausíveis), como por exemplo o estado 11133, em que o indivíduo tem muita dor e/ou desconforto, é extremamente ansioso e/ou deprimido, mas não tem nenhum tipo de problema nas demais dimensões de saúde.

4.3.1 Métricas de cálculo de inconsistências lógicas

Neste trabalho, analisamos a presença de inconsistências lógicas utilizando duas medidas, a proporção de respondentes inconsistentes na amostra e a taxa de média de inconsistências ponderada conforme proposto no estudo aprofundado das inconsistências lógicas realizado com a pesquisa aplicada ao estado de Minas Gerais (ANDRADE *et al.*, 2016).

A proporção de inconsistentes para cada método consiste na razão entre número de respondentes que cometeu pelo menos uma inconsistência lógica e o total de indivíduos em cada amostra, calculada para cada um dos métodos (ranking, EVA e TTO) em ambas as amostras (estado de Minas Gerais e amostra das capitais).

A taxa individual de inconsistências é calculada através da razão entre o número de inconsistências lógicas cometida por um indivíduo e o total de pares comparáveis que foram avaliados. A partir da taxa individual de inconsistência é possível calcular a taxa média de inconsistência na amostra. O intuito de calcular esta taxa é observar quantas das valorações produzidas violaram a estrutura lógica pré-definida dentre todos os pares comparáveis disponíveis ao entrevistado. O cálculo da taxa média de inconsistência requer que seja realizada ponderação da taxa individual de inconsistência pelo peso de cada indivíduo no total de pares comparados analisados (w_i). O peso de cada indivíduo pode diferir porque o número de estados de saúde comparáveis depende do grupo de estados de saúde avaliados. O peso (w_i) é definido pela razão entre o número de pares comparáveis para cada indivíduo (n_i) e a soma dos pares comparáveis de toda a amostra $\sum_{i=1}^N n_i$, ou seja:

$$W_i = \frac{n_i}{\sum_{i=1}^N n_i}$$

Esta média ponderada leva em conta a exposição relativa individual a cometer inconsistências lógicas.

A fim de ilustrar o cálculo da taxa individual de inconsistência apresentamos um exemplo hipotético. Suponha um indivíduo que produz o seguinte ordenamento entre quatro estados de saúde: (A) 11222 > (B) 11332 > (C) 11122 > (D) 33331. Esse indivíduo cometeu duas inconsistências lógicas: A > C e B > C dentre três pares comparáveis possíveis ([A,B]; [A,C] e [B,C]). Apesar de o estado de saúde D possuir 4 dimensões com graves problemas, não pode existir dominância lógica com os demais estados. Na comparação entre o par de estados A e D, por exemplo, verifica-se que nas quatro primeiras dimensões o estado D apresenta problemas graves enquanto o estado A apresenta problemas moderados ou não apresenta problemas. Na dimensão de ansiedade ou depressão o estado D não tem nenhum problema, enquanto que o estado A apresenta problemas moderados. Como não há uma relação de dominância em relação a esta dimensão, não está definida a comparabilidade entre este par de estados de saúde, o mesmo vale para os estados B e C. A taxa de inconsistência deste indivíduo é 2/3, 66,7%. A taxa de inconsistências individual varia de 0 (se nenhuma inconsistência lógica foi cometida entre todos os pares de estados de saúde comparáveis) até 1 (todos os pares comparáveis foram avaliados de forma inconsistente).

Um dos aspectos que pode afetar a presença de inconsistências se refere as características dos estados de saúde avaliados. Alguns estudos anteriores já evidenciaram que a distância entre os estados de saúde importa para explicar as inconsistências (ANDRADE *et al.*, 2016). Estados de saúde muito próximos e, portanto, com distancia inferior a 2, são mais difíceis de serem distinguidos pelos indivíduos, enquanto que estados de saúde mais distantes apresentam uma relação de dominância mais explícita.

Nesse trabalho, a análise das inconsistências lógicas também procurou controlar pela distância entre os pares comparáveis. Foram considerados os intervalos de distância de 1 a 2, 3 a 4, 5 ou mais estados de saúde. Conforme a distância entre os estados de saúde se eleva, espera-se que o

número de inconsistências cometidas se reduza, pois fica mais evidente a discrepância entre os estados de saúde.

4.4-Analise das Inconsistências segundo características individuais

Para analisar a associação entre características individuais e presença de inconsistências duas abordagens foram realizadas. Inicialmente foram testadas diferenças nas taxas medias de inconsistência observadas entre grupos específicos e sem seguida procedeu-se a analise controlada.

Modelos logísticos foram implementados para identificar se as características individuais poderiam ter algum impacto sobre a probabilidade de se cometer inconsistências. Outros trabalhos (BANSBACK, 2012; ANDRADE, 2013; CHEVALIER e POUVOURVILLE, 2013) já demonstraram que escolaridade pode ter efeitos sobre a capacidade de compreensão do instrumento e conseqüentemente uma menor frequência de respostas inconsistentes.

4.4.1 Modelo Logit

Seja a variável resposta y_i dicotômica, pode-se estimar uma variável latente y_i^* , tal que:

$$y_i^* = x_i\beta + u_i$$

Com uma regra de decisão tal que:

$$y_i = \begin{cases} 1 & \text{se } y_i^* > 0 \\ 0 & \text{se } y_i^* \leq 0 \end{cases}$$

Dessa forma, quanto maior y_i^* , maior deverá ser a possibilidade da variável y assumir o valor 1.

Tomando então:

$$P_i = Prob(y_i = 1) = P(y_i^* > 0) = P(x_i\beta + u_i > 0) = P(u_i > -x_i\beta)$$

O que implica que:

$$P_i = P(y_i = 0) = P(u_i \leq -x_i\beta)$$

Sendo $F(\cdot)$ a distribuição acumulada do erro (u), e que u e x sejam independentes, sendo que u tenha distribuição aproximadamente logística, temos:

$$P(y_i = 0) = F(-x_i\beta)$$

$$P(y_i = 1) = 1 - F(-x_i\beta) = F(x_i\beta)$$

Como u_i segue uma distribuição logística, temos que:

$$P(y_i = 1) = \frac{1}{1 + \exp(-x_i\beta)}$$

$$P(y_i = 0) = 1 - \frac{1}{1 + \exp(-x_i\beta)} = \frac{\exp(-x_i\beta)}{1 + \exp(-x_i\beta)}$$

$$\frac{P(y = 1)}{P(y = 0)} = \exp(-x_i\beta) \Rightarrow \ln\left(\frac{P(y = 1)}{P(y = 0)}\right) = x_i\beta$$

O logaritmo natural entre a probabilidade de sucesso e de fracasso é conhecido como razão de chance e revela quantas vezes algum evento é mais provável de acontecer do que não acontecer. Será sempre um número positivo, com limite inferior igual a zero. Caso o evento tenha probabilidade muito superior à de fracasso, terá um valor superior à unidade.

O vetor de parâmetros β é obtido através de uma estimação de máxima verossimilhança, permitindo a ocorrência de heteroscedasticidade na variância condicional de y dado x ($\text{Var}(y|x)$). Numa amostra aleatória, com tamanho n , a variável explicada y do tipo Bernoulli, a função de verossimilhança é dada por:

$$L(\beta) = \prod_{i=1}^n \{(P(y_i = 0))^{1-y_i} (P(y_i = 1))^{y_i}\}$$

Os estimadores de máxima verossimilhança do vetor β são funções de valores amostrais que maximizam $L(\beta)$ em relação a β (CAMERON e TRIVEDI, 2005). Nos modelos propostos, a variável resposta assume o sucesso quando o indivíduo comete ao menos uma inconsistência lógica, e zero caso não incorra em desvios na ordenação de dominância entre os estados de saúde.

4.4.2 Variáveis utilizadas no modelo.

Na regressão logística, a variável resposta dicotômica y_i assume valor igual um se indivíduo comete ao menos uma inconsistência lógica; e zero, caso o respondente seja totalmente consistente em suas decisões. A probabilidade de sucesso está associada a agentes inconsistentes. As variáveis independentes, listadas na TABELA 4, referem-se às características individuais dos respondentes, além de variáveis relativas ao procedimento de pesquisa.

Tabela 9. Variáveis utilizadas nos modelos Logísticos

Variáveis	Descrição
Variáveis Dependentes	
Modelo 0: Geral	Variável dicotômica, assume 1 se indivíduo cometeu alguma inconsistência, 0 caso contrário
Modelo 1: inconsistências com distância 1-2	Variável dicotômica, assume 1 se indivíduo cometeu alguma inconsistência entre pares comparáveis de estados com distância igual a 1 ou 2, 0 caso contrário
Modelo 2: inconsistências com distância 3-4	Variável dicotômica, assume 1 se indivíduo cometeu alguma inconsistência entre pares comparáveis de estados com distância igual a 3 ou 4, 0 caso contrário
Modelo 3: inconsistências com distância 5+	Variável dicotômica, assume 1 se indivíduo cometeu alguma inconsistência entre pares comparáveis de estados com distância igual a 5 ou maiores que 5, 0 caso contrário
Variável Independente	
Grau de escolaridade (incluída em todos os modelos)	Cinco variáveis para o nível de escolaridade individual <ul style="list-style-type: none"> 1. 1-3 anos de estudo (referência) 2. 4-7 anos de estudo 3. 8-10 anos de estudo 4. 11-14 anos de estudo 5. 15+ anos de estudo
Variáveis de controle	
Idade (incluída em todos os modelos)	Variável contínua, intervalo de 18 a 64 anos
Sexo (incluída em todos os modelos)	Variável dicotômica, sexo feminino como referência
Classe socio-econômica (incluída em todos os modelos)	Quatro variáveis para cada nível socioeconômico <ul style="list-style-type: none"> 1. Classe A (referência) 2. Classe B 3. Classe C 4. Classe D-E
Saúde auto-avaliada (incluída em todos os modelos)	Conjunto de 3 variáveis categóricas para saúde auto avaliada <ul style="list-style-type: none"> 1. Muito boa ou boa (referência) 2. Regular 3. Ruim ou Muito Ruim
Plano de Saúde privado (incluída em todos os modelos)	Variável dicotômica igual a 1 se indivíduo possui plano de saúde privado e 0 caso contrário
Entrevistador (incluída no modelo de Minas Gerais, mas não incluída na pesquisa das demais cidades)	Conjunto de 20 variáveis para cada entrevistador

Os critérios de riqueza se basearam no “Critério Brasil” (ABEP, 2013), que mede a riqueza domiciliar com base no grau de instrução do chefe de família e na posse de alguns itens domésticos para aferir o nível de renda, uma vez que esta é uma das variáveis econômicas com maior erro de medida. Existe uma função ponderada, que atribui valores numa escala de 0 a 46, subdividindo a população em 8 sub-amostras: A1, A2, B1, B2, C1, C2, D e E. Neste trabalho a subdivisão destes grupos foi sintetizada em 4 categorias.

A construção das variáveis dependentes do modelo baseou-se num critério de remoção também adotado por Andrade *et al.*, (2015). A lógica de realizar exclusões seletivas baseada no critério de distância entre pares de estados comparáveis é identificar se os desvios lógicos cometidos pelos respondentes se concentravam entre estados próximos ou se poderiam dar indícios da incapacidade de perceber a diferença entre estados muito discrepantes, cujas distâncias fossem maiores do que 5 na comparação entre os pares. Caso os desvios lógicos ocorram entre estados muito distintos, isto pode indicar que o respondente não compreendeu adequadamente a atividade proposta.

A primeira rodada de exclusões eliminou da amostra os indivíduos que cometeram ao menos uma inconsistência lógica entre pares comparáveis que estivessem em 5 ou mais graus de gravidade de distância. Nesta categoria, a diferença entre os estados é grande e deve ser mais facilmente notada, de modo que os indivíduos que não conseguiram perceber a diferença entre estes estados provavelmente não foram capazes de revelar suas preferências por estados de saúde de forma adequada. A segunda rodada de exclusões eliminou indivíduos que cometeram inconsistências entre pares de estados com distâncias entre 3 e 4. Neste grau ainda é possível perceber mais claramente a diferença entre estados de saúde, apesar de não ser tão explícita quando a primeira rodada de exclusões.

A última rodada de exclusões eliminou da amostra indivíduos que cometeram inconsistências lógicas entre pares de estados com distância de 1 a 2 graus de gravidade. Nesta categoria a distância entre os pares de estados é muito próxima, e por vezes as inconsistências lógicas acabam ocorrendo por desatenção do respondente ou incapacidade de perceber a distinção entre os estados.

4.5 – Estimação e comparação dos modelos de regressão para a EVA e TTO

A análise de regressão foi utilizada para estimar os valores para os parâmetros de valorização dos estados de saúde, baseando-se na mesma metodologia empregada por Andrade *et al.*, (2013). O objetivo destas estimações é estudar o impacto de exclusões de indivíduos que cometem inconsistências entre estados de saúde mais distantes e investigar os potenciais efeitos sobre a

qualidade do ajuste dos modelos. Foram realizadas estimações para o TTO e para a EVA, nas duas amostras, baseando-se nos mesmos critérios de estimação e criação de sub-amostras, a fim de realizar um estudo comparativo entre o desenho de estudo balanceado (Minas Gerais) e o estudo desbalanceado (amostra das capitais). As variáveis utilizadas nestas estimações estão descritas na tabela 10.

Os valores estimados para os estados de saúde, foram comparados com os valores diretamente observados. No caso do TTO, a saúde perfeita e a morte assumem os valores 1 e 0, respectivamente, em decorrência de serem os estados âncora deste método.

Os modelos de regressão de painel foram empregados, seguindo a mesma metodologia sugerida por Andrade *et al.* (2013). A utilização deste tipo de modelo é indicada pois cada indivíduo pode apresentar um padrão de respostas próprio, o que reforça a escolha por um modelo de efeitos aleatórios para controlar suas características individuais.

A escolha pelo modelo de efeitos aleatórios para a pesquisa aplicada em Minas Gerais (ANDRADE *et al.*, 2013) empregou os testes de Hausman (HAUSMAN; McFADDEN, 1984) e de Breusch-Pagan. O primeiro propõe uma comparação entre o modelo de efeitos fixos contra o modelo de efeitos aleatórios. Sua hipótese nula é que os efeitos individuais não são correlacionados com nenhuma outra covariada. O teste de Breusch-Pagan para heteroscedasticidade é feito sobre um modelo de Mínimos Quadrados Ordinários para testar se a variância é constante. Caso essa hipótese nula seja rejeitada, sugere-se o uso de um modelo de efeitos aleatórios que leve em conta a heteroscedasticidade nos resíduos.

As medidas de qualidade de ajuste dos modelos se basearam no Desvio Absoluto Médio (DAM) e no número de estados de saúde que apresenta o $DAM > 0,05$. As variáveis utilizadas nestes modelos seguem na tabela 10. As variáveis dependentes utilizadas nestes modelos passaram por transformações.

No caso do TTO, a saúde perfeita e a morte assumem os valores âncora de 1 e 0, os valores atribuídos aos estados melhores do que a morte são dados pela fórmula $V = \frac{x}{10}$, sendo x o período

de tempo vivido com saúde integral. Já os estados classificados como piores do que a morte, os valores são obtidos por $V = \frac{-x}{(10-x)}$. Nesta configuração, os valores obtidos no exercício variam de 1 até -19, o que pode enviesar negativamente os pesos obtidos no TTO. Uma correção para esta assimetria foi sugerida por Dolan *et al.*, (1996), que propõe que os valores negativos sejam ajustados pela fórmula $V_i = \frac{V}{(1-V)}$, se $V < 0$. Com esta adaptação, os estados piores do que a morte ficam limitados inferiormente a -1, criando um intervalo equidistante (-1 até 1), centrado no estado morte (0).

Já com relação à escala visual analógica, o ajuste feito visa ancorar os estados de saúde entre a morte (0) e a saúde perfeita (1). Para que seja possível a estimação dos parâmetros de valorização dos estados de saúde através da EVA faz-se necessária a adaptação dos valores pela fórmula:

$$Ve_j^t = \frac{EVA_j - EVA_{morte}}{EVA_{11111} - EVA_{morte}}$$

Tabela 10. Descrição das variáveis utilizadas nos modelos de estimação dos pesos para estados de saúde

Variáveis	Descrição
Variáveis Dependentes	
TTO	1-Vt
EVA	1-Ve
Variáveis Independentes	
MO2	igual a 1 se dimensão mobilidade for igual a 2
MO3	igual a 1 se dimensão mobilidade for igual a 3
SC2	igual a 1 se dimensão cuidados pessoais for igual a 2
SC3	igual a 1 se dimensão cuidados pessoais for igual a 3
UA2	igual a 1 se dimensão atividades usuais for igual a 2
UA3	igual a 1 se dimensão atividades usuais for igual a 3
PD2	igual a 1 se dimensão dor e desconforto for igual a 2
PD3	igual a 1 se dimensão dor e desconforto for igual a 3
AD2	igual a 1 se dimensão ansiedade ou depressão for igual a 2
AD3	igual a 1 se dimensão ansiedade ou depressão for igual a 3

Fonte: Elaboração do autor, adaptado de Andrade *et al* (2013);

Após realizadas as estimações, prossegue-se para a produção de sub-amostras com a exclusão seletiva de respondentes baseando-se nas distâncias em que as inconsistências lógicas foram

cometidas. Objetiva-se verificar se a qualidade do ajuste dos modelos estimados melhora com a exclusão seletiva. A primeira sub-amostra exclui respondentes que cometeram inconsistências com distância 8 ou mais, a segunda exclui indivíduos que cometeram inconsistência com 7 ou mais graus, e assim sucessivamente até a exclusão de todos os inconsistentes. São estimados os desvios absolutos médios para cada sub-amostra e a quantidade de n indivíduos é excluída em cada etapa.

5. Resultados

Este capítulo está dividido em 3 seções. Na primeira seção apresentamos a análise da proporção e taxa ponderada média de inconsistências, comparando os resultados obtidos nas pesquisas de Minas Gerais e das capitais. Na segunda subseção investigamos a associação entre as características individuais sobre a probabilidade de ocorrência de inconsistências lógicas dos indivíduos. Na terceira estimamos os modelos controlando pela distância entre os pares comparáveis de estados de saúde.

A maior parte das valorações inconsistentes ocorre quando estados de saúde mais próximos são comparados, sugerindo que a inconsistência nas valorações pode decorrer de dificuldade na percepção das diferenças na descrição dos estados de saúde. Quando estados de saúde muito distantes são avaliados de forma inconsistente, problemas na compreensão do exercício podem ter ocorrido, dificultando a aferição das preferências destes indivíduos. Na última seção do capítulo analisamos o efeito da exclusão seletiva de indivíduos que cometeram inconsistências na estimação das preferências sociais. Essa análise é importante para subsidiar a decisão de exclusão desses indivíduos da amostra

Nas análises aqui apresentadas serão feitas comparações entre os dois estudos a fim de investigar se o desbalanceamento dos estados de saúde importa para explicar a presença de valorações inconsistentes.

5.1. Métricas de inconsistências lógicas

A tabela 11 apresenta os resultados para a proporção de respondentes inconsistentes. Para ambos estudos os percentuais são próximos e evidenciam que cerca de 50% dos entrevistados comete ao menos uma inconsistência, independentemente do exercício de valoração (ranking, EVA ou TTO).

O exercício do ordenamento foi o método que apresentou maior proporção de indivíduos inconsistentes (53,13% nas capitais e 53,48% no estado). Este resultado pode ser atribuído à

poucafamiliaridade com a descrição dos estados de saúde, já que o ordenamento é o primeiro exercício de valoração. É neste exercício que os indivíduos compreendem a forma de descrição dos estados de saúde.

A escala visual analógica foi o método de valorização que apresentou as menores proporções de respondentes inconsistentes, 46,6% na pesquisa das capitais e 46,3% na pesquisa do estado. Esses resultados mostram que os indivíduos têm maior facilidade de realizar a valorização dos estados após a realização do ordenamento. É importante destacar que tanto no ranking quanto na EVA os indivíduos podem observar todos os estados de saúde disponíveis durante o exercício de valoração, de modo que a comparação entre eles é visual e direta.

No exercício do *Time Trade Off*, a proporção de inconsistentes foi maior do que a verificada no exercício da escala visual analógica. Outros estudos (BADIA, MONTSERRAT e HERDMAN, 1999; LAMERS *et al.*, 2006; SHAW, COONS e JOHNSON, 2002) também verificam essa relação entre o TTO e EVA. A amostra das capitais apresentou 51,01% indivíduos inconsistentes enquanto a amostra do Estado apresentou 51,69%. O maior percentual observado para o TTO em comparação à escala visual analógica provavelmente decorre da maior complexidade desse exercício. Embora o TTO seja o último exercício de valoração que o indivíduo realiza, o efeito aprendido observado entre os exercícios do Ranking e da EVA não compensa a maior a sua maior complexidade. No TTO os indivíduos não só devem entender a descrição dos estados de saúde e avaliá-la, como também estabelecer essa avaliação indiretamente através da troca entre tempo e qualidade de vida. Além disso, no TTO, não é possível aos indivíduos estabelecerem uma comparação direta entre estados de saúde, uma vez que a valoração ocorre em relação aos estados âncora, saúde perfeita ou morte. Cabe ainda mencionar que o exercício de valoração dos estados realizado no TTO ocorre sequencialmente, uma vez que os indivíduos valoram um estado de cada vez. Como o exercício é sequencial, as inconsistências podem também decorrer do efeito memória dado que os indivíduos podem não se lembrar das avaliações que fizeram em relação aos estados anteriormente avaliados, principalmente se estes estados forem muito próximos. Uma agenda futura de pesquisa pode investigar em que medida a ordem de apresentação dos estados de saúde impacta na chance de realização de valoração inconsistente.

A análise da proporção de indivíduos inconsistentes sugere que o balanceamento do conjunto de estados de saúde apresentados aos indivíduos para valoração não altera a quantidade total de indivíduos que comete ao menos um desvio lógico. Esse resultado independe do exercício de valoração realizado. Entre os três métodos, a EVA foi a que apresentou a menor proporção de indivíduos inconsistentes nos dois estudos.

Tabela 11. Proporção e Taxas de Inconsistências individuais para ambas as amostras

	Brasil		
	Rank	EAV	TTO
% respondentes com inconsistências lógicas %(IC 95%)	53,13 [51,84; 54,42]	46,60 [45,31; 47,89]	51,05 [49,76; 52,35]
Taxa de Inconsistência Lógica			
Média	14,54% [14,04%; 15,03%]	12,18% [11,70%; 12,66%]	12,26% [11,80%; 12,71%]
Mediana	9,09%	0,00%	7,69%
	Minas Gerais		
% respondentes com inconsistências lógicas %(CI 95%)	53,48 [51,79; 55,17]	46,27 [44,58; 47,96]	51,69 [49,94; 53,38]
Taxa de Inconsistência Lógica			
Média	8,73% [8,35%; 9,15%]	7,26% [6,88%; 7,64%]	8,64% [8,23%; 9,05%]
Mediana	6,66%	0,00%	6,66%

FONTE: elaboração do autor

Com relação às taxas de inconsistência médias ponderadas, o resultado para os três exercícios de valoração entre as duas pesquisas apresentou diferenças importantes. Para os três exercícios, a taxa média de inconsistências lógicas foi mais elevada na amostra das capitais. O ordenamento apresentou a maior taxa (14,54%), seguido pela EVA (12,18%) e TTO (12,26%). No estado de Minas Gerais o patamar das taxas médias foi inferior ao observado para as capitais: 8,73% no Ranking, 8,48% no TTO e 7,26% na EVA. A tabela 12 apresenta os testes para igualdade entre as taxas de inconsistência média ponderadas, deve-se rejeitar a hipótese nula de igualdade entre as taxas das capitais e do estado para todos os exercícios a 1% de significância.

Tabela 12. Teste para igualdade nas taxas de inconsistência média ponderadas

Amostra	n	Média	Desvio Padrão	t-cal	t-critico	gl	p	decisão
TTO								
Capitais	5365	12,26	17,07	10,55	2,58	8546	0,00	Rejeição
Estado de Minas Gerais	3183	8,64	11,77					
EVA								
Capitais	5746	12,16	18,83	13,76	2,58	9089	0,00	Rejeição
Estado de Minas Gerais	3345	7,25	11,62					
Rank								
Capitais	5746	14,47	19,42	15,55	2,58	9089	0,00	Rejeição
Estado de Minas Gerais	3345	8,71	11,82					

Fonte: elaboração do autor, n: número de observações, t-cal: estatística T calculada, p: p-valor relativo à estatística T

As taxas médias de inconsistência encontradas na pesquisa realizada nas capitais destoam dos resultados de outros trabalhos. De modo geral, as taxas médias de inconsistência encontradas nos dois estudos brasileiros são mais elevadas do que aquelas estimadas em outros países. Na Espanha as taxas de inconsistência foram de 4% no TTO, 1,32% na EVA e 1,14% no ordenamento (BADIA, HERDMANN e MONTSERRAT; 1999); na Holanda, as taxas foram de 2% para a EVA e 4,3% para o TTO; na Dinamarca, as taxas de inconsistência foram de 4,28% para o ordenamento, 4,34% para a EVA e 7,33% para o TTO. Nesses países, o número de estados de saúde avaliados pelos indivíduos era maior (19 na Holanda e 13 na Espanha, 14 a 16 na Dinamarca), o que pode explicar a menor taxa observada, uma vez que pode diluir o número de inconsistências lógicas ao elevar o número de pares comparáveis possíveis.

A análise da taxa média de inconsistência ponderada sugere que o balanceamento dos estados de saúde parece afetar a realização de valores inconsistentes nos três exercícios. As taxas observadas na pesquisa das capitais são mais elevadas. A presença de estados leves, moderados e graves na amostra de Minas Gerais parece facilitar a percepção de diferença entre os estados de saúde.

5.2 Associação entre características Individuais e a taxa de inconsistência

A ocorrência das inconsistências lógicas pode estar associada a algumas características individuais. Para testar estas associações, testes de diferenças de média e modelos de regressão logística foram realizados a fim de identificar quais das características socioeconômicas ou demográficas podem estar associadas a valorações inconsistentes.

A tabela 13 apresenta a taxa média de inconsistência lógica para cada método de valoração, em ambas as amostras, de acordo com as características socioeconômicas e demográficas dos indivíduos. Maiores níveis de escolaridade estão associados com menores taxas de inconsistência de forma monotônica para todos os métodos, sendo esta diferença significativa a 5%¹ tanto na amostra de Minas Gerais quanto na amostra das capitais. A posse de plano de saúde privado e classes de renda mais elevadas também estão associadas a menores taxas de inconsistências, nos três métodos em ambas as amostras. Na pesquisa das capitais, indivíduos que auto avaliaram sua saúde como boa ou muito boa apresentaram menores as taxas de inconsistência lógicas, nos três exercícios (estatisticamente significante a 5%). Na pesquisa de Minas Gerais esse resultado também se verificou para o ranking e para a EVA.

Dolan (1996) relata que indivíduos que já tiveram experiências com doenças podem apresentar preferências por estados de saúde diferentes do resto da população. Sacklet e Torrance (1978) mostram também que pacientes que realizam hemodiálise costumam atribuir maiores valores para estados de saúde com maiores graus de debilidade (atribuem uma maior qualidade de vida para casos de hemodiálise dos rins do que o público em geral). Por outro lado, Llewellyn e Thomaset al. (1984), encontraram que a condição de saúde do paciente não interfere na forma como avalia os estados de saúde. No escopo deste trabalho a saúde auto avaliada será testada posteriormente em modelos logísticos para identificar se existe um efeito sobre a inconsistência dos indivíduos.

¹ Valores obtidos através do teste de Kruskal-Wallis.

Tabela 13. Taxa média de inconsistência lógica por características sócio demográficas e de saúde

Características Individuais	Minas Gerais			Brasil			
	Rank	EAV	TTO	Rank	EAV	TTO	
Genero	Homem	8.80%	7.41%	8.43%	14.18%	12.33%	16.65%
	Mulher	8.62%	7.07%	8.55%	14.80%	11.96%	15.36%
		<i>P-valor= 0.783</i>	<i>P-valor= 0.423</i>	<i>P-valor= 0.834</i>	<i>P-valor= 0.6807</i>	<i>P-valor= 0.455</i>	<i>P-valor= 0.0709</i>
Faixa etária	18-25	8.66%	6.84%	8.79%	13.62%	11.03%	15.61%
	26-35	9.03%	8.08%	8.55%	13.50%	11.38%	15.95%
	36-45	9.13%	7.47%	8.34%	16.30%	13.31%	16.00%
	46-55	8.65%	6.78%	8.46%	14.45%	12.29%	16.23%
	56-64	7.43%	6.75%	7.90%	14.59%	13.29%	16.80%
		<i>P-valor= 0.477</i>	<i>P-valor= 0.223</i>	<i>P-valor= 0.746</i>	<i>P-valor= 0.0055</i>	<i>P-valor= 0.011</i>	<i>P-valor=0.9784</i>
Nível Educacional	Prim. Inc.	10.11%	6.45%	11.03%	17.40%	13.28%	25.30%
	Prim. Comp.	9.41%	8.43%	9.22%	18.66%	17.38%	20.48%
	Fund. Comp.	10.06%	7.83%	9.85%	15.32%	12.01%	16.02%
	E.M. Comp.	7.93%	6.58%	7.49%	9.55%	7.44%	11.00%
	Sup. Comp.	5.50%	6.00%	5.61%	5.38%	4.18%	8.89%
		<i>P-valor<0.001</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>
Estado Civil	Casado	8.61%	7.51%	8.22%	14.39%	12.26%	15.84%
	Viúvo	10.41%	6.41%	11.99%	17.87%	16.37%	18.73%
	Divorciado	8.81%	7.33%	7.36%	14.38%	11.81%	16.19%
	Solteiro	8.72%	6.93%	8.84%	14.05%	11.60%	15.82%
		<i>P-valor= 0.461</i>	<i>P-valor= 0.526</i>	<i>P-valor= 0.029</i>	<i>P-valor= 0.1066</i>	<i>P-valor= 0.031</i>	<i>P-valor= 0.719</i>
Plano de Saúde	Sim	9.32%	6.40%	8.92%	11.50%	9.29%	13.01%
	Não	7.46%	7.66%	7.61%	16.80%	14.44%	18.49%
		<i>P-valor< 0.001</i>	<i>P-valor=0.012</i>	<i>P-valor=0.015</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>
Saúde Auto-avaliada	Muito Bom	7.73%	5.87%	7.96%	11.24%	9.12%	12.99%
	Bom	8.57%	7.33%	8.46%	14.14%	11.06%	15.64%
	Regular	9.98%	8.85%	9.25%	16.73%	15.74%	18.47%
	Ruim/Muito Ruim	14.39%	9.52%	9.83%	18.89%	15.85%	19.55%
		<i>P-valor= 0.006</i>	<i>P-valor<0.001</i>	<i>P-valor=0.219</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>	<i>P-valor<0.001</i>
Classe de Riqueza	A	5.21%	3.85%	6.72%	8.39%	6.93%	10.89%
	B	7.54%	6.42%	7.37%	12.95%	9.76%	14.12%
	C	9.58%	7.99%	9.12%	17.19%	15.62%	18.63%
	D/E	11.45%	8.69%	11.74%	15.92%	15.94%	24.94%
		<i>P-valor< 0.001</i>					

FONTE: elaboração do autor

Teste de Kruskal-Wallis para diferenças de médias, o nível de significância adotado foi de 5%

Com base nestes resultados, modelos logísticos gerais foram estimados para a amostra total, considerando também uma variável identificadora para as informações do estado de Minas Gerais. O objetivo destes modelos é mapear quais características individuais têm os maiores efeitos sobre a probabilidade de se cometer uma inconsistência lógica, os resultados seguem reportados na tabela 14.

Tabela 14. Resultados para o Modelo Geral Logístico Multivariado – Ranking, EVA e TTO- Amostra das Capitais e de Minas Gerais

	Ranking		EVA		TTO	
	Odds Ratio Ajustado	P Valor	Odds Ratio Ajustado	P Valor	Odds Ratio Ajustado	P Valor
Idade	1,001	0,495	1,002	0,278	0,995	0,010
Sexo (Categoria Ref.: Feminino)	1,037	0,390	0,984	0,709	0,934	0,126
Escolaridade (Categoria Ref.: 1-3 anos de estudo)						
4-7 anos de estudo	1,021	0,866	1,413	0,006	0,756	0,032
8-10 anos de estudo	0,863	0,258	1,147	0,285	0,619	0,000
11-14 anos de estudo	0,560	0,000	0,786	0,087	0,438	0,000
15+ anos de estudo	0,315	0,000	0,453	0,000	0,375	0,000
Significância conj. (prob > chi²)	0,000		0,000		0,000	
Classe Socioeconômica (Categoria Ref.: Class. A)						
Classe B	1,336	0,005	1,143	0,203	0,984	0,876
Classe C	1,503	0,000	1,382	0,004	0,989	0,920
Classes D e E	1,470	0,006	1,440	0,010	1,186	0,233
Significância conj. (prob > chi²)	0,002		0,000		0,319	
SaúdeAuto avaliada (Categoria Ref.: Muito bom ou bom)						
Regular	1,210	0,000	1,380	0,000	1,121	0,033
Ruim, Muito Ruim	1,180	0,226	1,011	0,929	1,339	0,039
Significância conj. (prob > chi²)	0,001		0,000		0,020	
Plano de Saúde privado	1,087	0,087	1,112	0,031	1,116	0,028
Minas Gerais	0,819	0,000	0,847	0,000	0,891	0,017
Número de obs.	9079		9079		8536	
R² ajustado	0,0244		0,0275		0,0142	

Nota: Todos os modelos incluíram constante (resultados omitidos na tabela)

Fonte: elaboração do autor

Os resultados representados na tabela 14 apontam que algumas características individuais possuem relação com valorações inconsistentes. A variável dependente dos modelos aqui estimados é uma variável latente para a probabilidade do indivíduo *i* ser inconsistente no exercício em questão. As variáveis independentes são características individuais coletadas na pesquisa: idade, sexo, nível de escolaridade, classe de riqueza (aproximada de acordo com o critério Brasil), saúde auto avaliada, posse de plano de saúde privado e uma variável

identificadora para o Estado de Minas Gerais. O objetivo da inclusão da variável dicotômica para o estado é verificar se esta amostra apresentou um comportamento diferente daquele observado nas capitais.

A probabilidade de ocorrência de inconsistências lógicas no exercício do ranking está associada de forma estatisticamente significativa com as variáveis de grau de escolaridade, classe de riqueza, saúde auto avaliada e se os indivíduos residiam no estado de Minas Gerais. Indivíduos residentes no estado de Minas Gerais possuem 19% a menos de chance de ser inconsistente do que o restante da amostra. Observa-se significância conjunta para a escolaridade neste exercício, maiores graus de instrução estão relacionados a menores probabilidades de ocorrência de valorações inconsistentes (um indivíduo com 11-14 anos de escolaridade tem 44% menos chances de ser inconsistente do que um indivíduo com 1-4 anos de estudo, para indivíduos com 15 ou mais anos de estudo, essa probabilidade é 68,5% menor). A relação entre a riqueza e a probabilidade de inconsistência é negativa, quanto mais ricos, menores as chances de ser inconsistentes (um indivíduo das classes D ou E tem 47% a mais de chance de cometer ao menos um desvio lógico do que um indivíduo pertencente à classe A). Indivíduos que avaliam a própria saúde como regular, ruim ou muito ruim possuem maiores chances de cometer valorações inconsistentes do que respondentes que se auto avaliam saudáveis. As variáveis de sexo, idade e posse de plano de saúde não apresentaram significância estatística para explicar a probabilidade de inconsistência.

As características individuais que estão relacionadas à probabilidade de inconsistência para a Escala Visual Analógica são as mesmas que afetam as valorações do ranking. A relação entre a chance de ser inconsistente e o grau de escolaridade é inversa, um indivíduo com 15 ou mais anos de estudo apresenta 54,7% menos chances de ser inconsistentes na EVA do que um indivíduo com 1 a 4 anos de escolaridade. Quanto menor a classe de renda do indivíduo, maiores as probabilidades do mesmo ser inconsistente (sujeitos pertencentes à classe de renda D/E têm 44% mais chances de ser inconsistentes do que sujeitos da classe A). Indivíduos que se auto avaliam com saúde regular são mais propensos a cometer inconsistências do que indivíduos que consideram a própria saúde como boa ou muito boa (38% a mais de probabilidade). Entrevistados que não possuíam plano de saúde privado apresentaram maiores chances de ser inconsistentes

(11,2%) do que aqueles que possuem e os residentes de Minas Gerais foram menos propensos a cometer inconsistências lógicas do que o restante da amostra (15,3% a menos de chance de cometer uma inconsistência).

Para o *Time Trade Off* a relação entre as características individuais e a razão de chance de inconsistências lógicas foi similar aos outros dois exercícios. Destaca-se o papel da escolaridade, já que um indivíduo com 15 ou mais anos de estudo apresentou 62,5% a menos de chance de cometer uma inconsistência do que um indivíduo com escolaridade de 1-4 anos. Indivíduos que não se avaliaram com saúde muito boa ou boa também foram mais propensos a cometer desvios lógicos neste exercício, assim como aqueles que não possuem plano de saúde. Os residentes de Minas Gerais tiveram 10,9% a menos de chance de ser inconsistentes do que os residentes das capitais. Neste exercício a variável de riqueza não apresentou significância estatística para explicar variações na probabilidade de inconsistência individual.

De maneira geral, a ocorrência de inconsistências lógicas está associada ao grau de escolaridade dos indivíduos, quanto mais escolarizado, maior a familiaridade com pesquisas científicas e o seu propósito. Adicionalmente, por se tratar de um processo longo de entrevista, os indivíduos que compreendem com maior facilidade o que está sendo proposto têm menores chances de apresentar cansaço ao longo do processo, o que pode ajudar a diminuir o número de valorações inconsistentes. Craig e Ramachandran (2006) apontam também que indivíduos que apresentam enfado ou se cansam da pesquisa podem se tornar “sabotadores” (termo de cunho dos autores), isto é, deliberadamente fornecer respostas que fogem do comportamento racional esperado e que não refletem suas preferências pelos estados de saúde. Seu único objetivo passa a ser antecipar o final da entrevista. O estrato social que comete um menor número de inconsistências lógicas tende a ser mais desejável para realizar a coleta de dados, uma vez que revela com maior clareza as suas preferências, contudo, como se trata de uma pesquisa voltada para subsidiar a tomada de decisão para intervenções de saúde pública, deve-se priorizar a coleta das preferências de todas as camadas sociais.

A saúde auto avaliada apresentou significância estatística nos três exercícios, de forma similar. Indivíduos que consideram a própria saúde como regular, ruim ou muito ruim são mais propensos

a incorrer em desvios lógicos do que indivíduos saudáveis. Este resultado deve ser interpretado com maior profundidade, uma vez que estes respondentes podem apresentar um perfil de preferências distinto dos demais. De acordo com Dolan (1996), pessoas que sofrem com alguma condição de saúde mais severa podem atribuir valorações mais altas para estados de saúde mais debilitados, de modo que a ocorrência do desvio da ordenação lógica é decorrente de preferências individuais mais amenas em relação a estados de saúde com debilidades mais agudas.

A relação que existe entre a ocorrência de inconsistências e a classe de renda verificada nos exercícios do ordenamento e da escala visual analógica não se verifica no caso do *Time Trade Off*. O efeito verificado nos dois primeiros exercícios é uma relação inversa entre riqueza e a ocorrência de inconsistências, quanto mais rico, menores as chances de desvios lógicos. Este efeito também pode se estender à posse de plano de saúde privado, já que indivíduos que não possuem plano também apresentaram maiores probabilidades de ser inconsistentes.

Nos três exercícios, indivíduos que são do estado de Minas Gerais apresentaram menores probabilidades de ser inconsistentes. Este efeito pode indicar que a escolha pela metodologia de blocos de estados de saúde fixos balanceados tem a capacidade de reduzir a incidência de inconsistências lógicas na pesquisa. Este efeito também pode ser verificado nas taxas de inconsistências médias ponderadas mais baixas verificadas nesta população. Entretanto, por se tratar de uma pesquisa para obtenção de preferências sociais, essas diferenças podem estar associadas a diferentes preferências entre indivíduos residentes nas três cidades e indivíduos residentes em Minas Gerais. As informações disponíveis não permitem separar esses dois elementos que podem afetar a presença de inconsistências lógicas. Por se tratar de inconsistências, nossa hipótese é que essas diferenças são mais fortemente associadas as diferenças no desenho do estudo, o ou seja, ao desbalanceamento dos estados de saúde avaliados. Este trabalho é uma oportunidade singular de realizar estas comparações, tendo em vista a estrutura que abrange três regiões do país, seguindo um protocolo padronizado, com diferenças metodológicas na escolha dos estados de saúde apresentados aos indivíduos.

O modelo logístico apresentado para o ranking apresenta um R^2 de 0,024 com 9054 observações. A qualidade do ajuste foi relativamente baixa, o que também ocorreu para o modelo da Escala

Visual Analógica e para o TTO (com R^2 de 0,0275 e 0,0142, respectivamente). Este resultado dá indícios de que outras variáveis também podem estar relacionadas a ocorrência de inconsistências lógicas. O objetivo central destes modelos não é buscar um elevado grau de ajuste, mas identificar a possível relação entre as características individuais e a probabilidade de ocorrência de inconsistências lógicas individuais.

Com base no exposto, as características individuais que afetam a probabilidade de o indivíduo cometer um desvio lógico são aparentemente as mesmas nos três exercícios de obtenção de preferências. A ocorrência das valorações inconsistentes parece estar associada a níveis de escolaridade mais baixos, com níveis de renda inferiores. A análise a seguir busca isolar o efeito da distância entre os estados de saúde, visando refinar os critérios de análise das inconsistências lógicas.

5.3 Relação entre inconsistências lógicas e as distâncias dos pares de estados de saúde

A investigação proposta nesta seção aborda as inconsistências lógicas levando em consideração a distância entre os pares comparáveis. Na Figura 5 pode-se observar que em ambas as pesquisas a proporção de respostas inconsistentes é inversamente relacionada com a distância entre os estados de saúde, de forma monotônica, quanto maior a distância, menor a frequência de inconsistências nas valorações individuais.

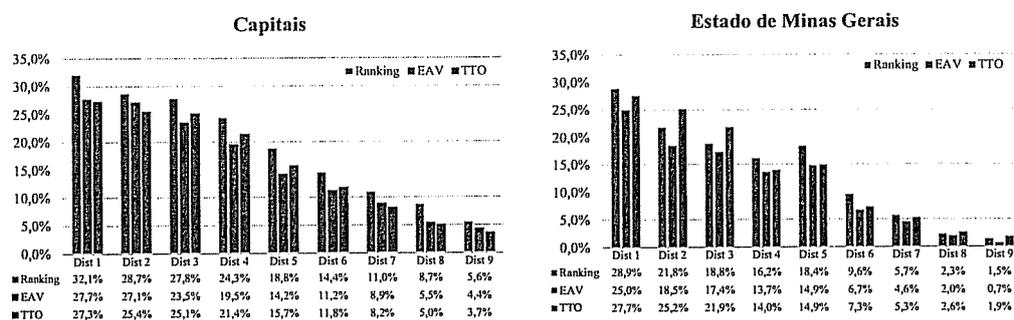


FIGURA 5. Proporção de Inconsistências Lógicas por distância para o Ranking, EVA e Time Trade Off, pesquisa das capitais e do Estado de Minas Gerais

FONTE: elaboração do autor

A figura 5 evidencia a relação inversa entre a proporção de inconsistências com a distância entre os pares comparáveis analisados. Quanto maior a distância entre os pares, menor a proporção de desvios lógicos, em ambas as pesquisas e nos três métodos. Ocorre um decaimento gradativo da proporção de respostas inconsistentes conforme a distância dos pares se eleva. A amplitude deste intervalo é grande: na pesquisa nacional, varia entre 32,1% até 5,6% no ranking; 27,7% até 4,4% para a escala analógica; 27,3% até 3,7% para o Time Trade Off. Em Minas Gerais o intervalo de variação foi de 28,9% até 1,5% no ranking; 25,0% até 0,7% na EVA e 27,7% até 1,9% no TTO. Pode-se ver que as proporções de inconsistências são similares para pares comparáveis com menores distâncias nas duas amostras, mas que a pesquisa das capitais apresenta percentuais mais elevados nas maiores distâncias. Essa diferença pode ser decorrente do desenho de estudo, já que a escolha balanceada de estados de saúde utilizada em Minas Gerais propõe uma distinção mais clara entre eles.

Os indivíduos apresentaram menores percentuais de repostas inconsistentes para todas as distâncias na Escala Visual Analógica em ambas as pesquisas. Este resultado sugere que existe um efeito de aprendizado interno da pesquisa, já que a EVA sucede o ordenamento no protocolo de aplicação, e apresenta menores proporções de inconsistência. O TTO apresenta maiores proporções do que a escala analógica pois é um método que não permite comparações diretas com os demais estados. Na figura 2, é feita análise sobre as taxas médias de inconsistência, levando em conta as distâncias.

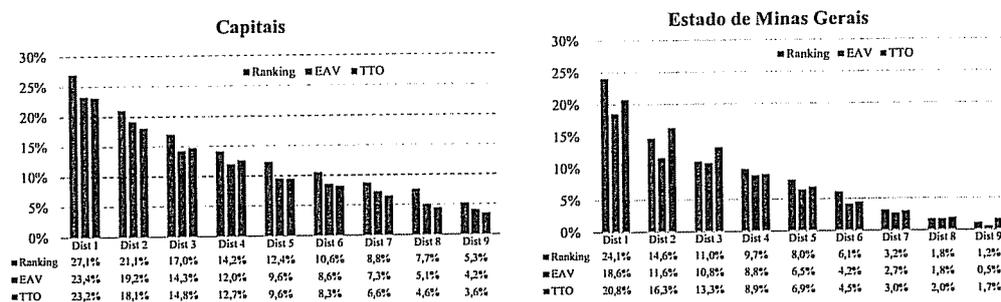


FIGURA 6. Média ponderada da taxa de inconsistência para o ranking, EVA e TTO, pesquisa das capitais e do Estado de Minas Gerais

FONTE: elaboração do autor

Nas taxas de inconsistências lógicas o efeito do balanceamento do estudo tem impactos mais pronunciados. Aqui também se observa um decaimento das taxas conforme a distância entre os pares de estados se eleva, mas este fenômeno ocorre com maior intensidade na pesquisa realizada no estado de Minas Gerais. Atribuímos este efeito ao balanceamento de estados de saúde, já que a presença de estados graves, moderados e leves pode facilitar a comparação entre os pares de estados de saúde.

A amplitude de variação entre as taxas de inconsistência também foi grande, abrangendo 27,1% até 5,3% no ranking, 23,4% até 4,2% na EVA e 23,2% até 3,63% no TTO para a pesquisa nacional. Na pesquisa de Minas Gerais o intervalo de variação foi de 24,1% até 1,2% para o ranking; 18,6% até 0,6% para a EVA e 20,8% até 1,7% no TTO. Pode-se ver que as taxas de inconsistências são mais baixas na pesquisa de Minas Gerais em todos os exercícios, com um decaimento maior do que aquele verificado na outra pesquisa. As taxas de inconsistência médias para pares de estados com maiores distâncias ficam mais próximas a 0 no estado de Minas Gerais, enquanto que nas capitais são superiores a 3,6%.

5.3.1 Modelos logísticos considerando as distâncias:

Os modelos apresentados nas tabelas 15 são propostos para investigar como as características individuais afetam a probabilidade de cometer inconsistências, controlando pelas distâncias entre os pares comparáveis. A variável dependente dicotômica do modelo 1 assume valor 1 caso alguma inconsistência entre pares de distância 1 ou 2 tenha ocorrido e 0 caso o indivíduo seja plenamente consistente. A variável dependente dicotômica do modelo 2 é igual a 1 caso alguma inconsistência tenha ocorrido em pares de estados com distâncias de 3 ou 4 e 0 caso contrário. A variável dependente dicotômica do modelo 3 assume 1 caso alguma inconsistência tenha ocorrido entre pares de estados de saúde com distâncias iguais ou superiores a 5 e 0 caso contrário. O objetivo é realizar a comparação entre os resultados obtidos nas duas pesquisas, seguindo especificação similar àquela apresentada na tabela 13. Para investigar da forma mais isolada possível os efeitos do balanceamento dos estados de saúde aqui também será empregada uma variável identificadora para os entrevistados no estado de Minas Gerais.

Uma consideração inicial importante é o número de observações utilizado em cada modelo, nos três exercícios o número de observações do modelo 1 é inferior ao dos modelos 2 e 3. Isso reflete que muitos indivíduos da amostra não se depararam com comparações entre estados de saúde com distância 1 ou 2.

Os resultados mostram que a escolaridade afeta a probabilidade de inconsistências mesmo levando em consideração a distância, apenas o modelo 1 do TTO não apresentou significância conjunta para a escolaridade, mas os modelos 2 e 3 mostram que indivíduos com mais anos de estudos tem chances mais baixas de incorrer em desvios lógicos. Este resultado tem especial relevância pois indica que mesmo para respondentes muito educados, a distinção entre estados de saúde muito próximos pode ser complexa e muitas vezes este tipo de comparação pode fazer com que o indivíduo não seja capaz de perceber a distância entre os estados de saúde.

A classe de riqueza teve significância conjunta nos modelos da Escala Visual analógica, da mesma forma dos modelos gerais, apresentados na tabela 13. Para o exercício do ranking, observa-se que apenas no modelo 3 existe significância conjunta desta variável. Este resultado demonstra que é importante o refinamento da análise, levando em conta as distâncias das inconsistências. Indivíduos com menores condições materiais apresentam maiores chances de incorrer em desvios lógicos do que os indivíduos mais ricos na EVA. Nenhum modelo do TTO apresentou significância estatística para a variável de classe de riqueza.

A saúde auto avaliada foi conjuntamente significativa para a EVA e para o Ranking, sendo que indivíduos que se avaliam com saúde regular têm maiores probabilidades de incorrer em desvios lógicos. No *Time Trade Off* significância conjunta não pode ser verificada a 5% em nenhum dos modelos. A idade afetou a probabilidade de inconsistências lógicas para distâncias 1-2 e 3-4 no TTO, indivíduos mais velhos apresentaram menores chances de incorrer em inconsistências lógicas.

Para a pesquisa do estado de Minas Gerais, Andrade *et al.*, (2015) destacam que os níveis de escolaridade são conjuntamente significantes para o ranking e para o TTO, sendo que em ambos os métodos, quanto maior a escolaridade, menor a probabilidade de ser inconsistente. Na EVA, a

classe socioeconômica (modelos 1 e 3) e saúde auto avaliada (modelo 3) são conjuntamente significantes. Além disso ressaltam a significância conjunta para a variável de entrevistador, independentemente do exercício utilizado (resultados não mostrados na tabela). Não foi possível reproduzir estes resultados na amostra completa.

Os indivíduos residentes no Estado de Minas Gerais apresentaram menores probabilidades de incorrer em inconsistências lógicas em todos os exercícios, para todas as distâncias consideradas. Este resultado corrobora a qualidade do estudo realizado em Minas Gerais, pois mesmo utilizando uma quantidade menor de estados de saúde na pesquisa, pode-se chegar em níveis mais baixos de inconsistências lógicas do que na pesquisa das capitais, mesmo levando em conta a distância entre os pares de estados de saúde comparáveis.

A escolaridade é a característica com maior impacto sobre a consistência dos indivíduos, mesmo levando em conta a distância entre os pares de estados. O efeito do balanceamento dos estados de saúde afeta de forma significativa a taxa de inconsistência individual, mas os atributos individuais que explicam este fenômeno aparentemente são os mesmos em ambas as pesquisas.

Tabela 15. Resultados para o Modelo Geral Logístico Multivariado – Ranking, EVA e TTO- Amostra das Capitais e de Minas Gerais

Variáveis	Rank						EVA						TTO					
	Modelo 1		Modelo 2		Modelo 3		Modelo 1		Modelo 2		Modelo 3		Modelo 1		Modelo 2		Modelo 3	
	OR	P Valor																
Idade	0,997	0,261	1,001	0,074	1,001	0,150	1,000	0,901	1,000	0,089	1,002	0,391	0,994	0,030	0,996	0,029	0,998	0,263
Sexo (Cat. ref.: Feminino)	1,022	0,739	1,019	0,303	1,026	0,684	1,077	0,269	0,949	0,305	0,959	0,458	0,953	0,481	0,919	0,090	0,976	0,666
Escolaridade (Categoria de ref.: 1-3 anos de estudo)																		
4-7 anos de estudo	0,704	0,058	0,941	0,946	1,127	0,455	1,153	0,454	1,427	0,013	1,705	0,001	0,865	0,450	0,781	0,064	0,788	0,086
8-10 anos de estudo	0,697	0,058	0,767	0,226	0,962	0,928	1,027	0,891	1,080	0,602	1,234	0,222	0,827	0,335	0,613	0,000	0,584	0,000
11-14 anos de estudo	0,521	0,002	0,514	0,001	0,574	0,002	0,853	0,461	0,762	0,102	0,859	0,436	0,760	0,203	0,444	0,000	0,372	0,000
15+ anos de estudo	0,288	0,000	0,320	0,000	0,321	0,000	0,562	0,025	0,376	0,000	0,355	0,001	0,705	0,186	0,394	0,000	0,218	0,000
Significância conj.	0,000		0,000		0,000		0,004		0,000		0,000		0,652		0,000		0,000	
Classes Socioeconômicas (Categoria de ref.: Classe A)																		
Classe B	1,349	0,067	1,141	0,570	1,565	0,039	1,192	0,286	1,041	0,762	1,178	0,337	0,842	0,267	1,017	0,888	1,101	0,967
Classe C	1,331	0,095	1,305	0,222	1,658	0,030	1,393	0,056	1,309	0,054	1,458	0,031	0,905	0,543	1,079	0,556	1,066	0,681
Classes D e E	1,377	0,138	1,266	0,484	1,932	0,713	1,669	0,019	1,214	0,253	1,516	0,042	1,196	0,403	1,220	0,216	1,244	0,241
Significância conj.	0,332		0,059		0,005		0,040		0,001		0,005		0,110		0,405		0,363	
Saúde Auto Avaliada (Categoria de ref.: Muito bom, bom)																		
Regular	1,254	0,004	1,135	0,032	1,141	0,114	1,333	0,000	1,310	0,000	1,376	0,000	1,096	0,270	1,103	0,098	1,096	0,164
Ruim, Muito Ruim	1,086	0,704	1,427	0,064	1,485	0,064	1,053	0,812	1,199	0,235	1,507	0,009	1,253	0,295	1,348	0,047	1,150	0,398
Significância conj.	0,017		0,010		0,006		0,001		0,000		0,000		0,362		0,054		0,310	
Plano de saúde privado	1,182	0,026	1,144	0,013	1,164	0,105	1,042	0,591	1,110	0,074	1,213	0,004	1,143	0,091	0,999	0,991	1,118	0,086
Estado de Minas Gerais	0,768	0,000	0,537	0,013	0,763	0,105	0,769	0,000	0,624	0,000	0,828	0,002	0,950	0,482	0,672	0,000	0,718	0,000
R ² ajustado	0,017		0,030		0,024		0,016		0,030		0,033		0,006		0,105		0,022	
Número de obs.	3974		8005		8434		3974		8005		8434		3747		7538		7924	

Nota: Todos os modelos incluíram as constantes (resultados omitidos na tabela)

5.3.2 Qualidade de ajuste de Modelos realizando exclusões seletivas baseadas na distância entre pares de estados de saúde.

Nesta seção serão apresentadas comparações da qualidade do ajuste de modelos de regressão realizando exclusões seletivas de indivíduos que cometem inconsistências lógicas na EVA e no Time Trade Off. O objetivo destas estimações é medir como a exclusão seletiva de indivíduos afeta a qualidade de ajuste do modelo e se o diferente desenho do estudo entre as pesquisas tem efeitos sobre estes resultados. Serão realizadas comparações entre os resultados verificados por Andrade *et al.*, (2015) com aqueles estimados com a pesquisa realizada nas capitais.

Os modelos de regressão aqui implementados visam a estimação dos pesos sociais por estados de saúde, a metodologia de estimação implementada utilizou regressões de painel com efeitos aleatórios. Foram utilizadas variáveis dicotômicas de identificação para cada grau de gravidade, conforme discutido na seção de aspectos metodológicos gerais. Foram estimados modelos para as amostras do estado de Minas Gerais e para a amostra das capitais. Serão realizadas exclusões seletivas de indivíduos que cometem inconsistências lógicas para avaliar como a qualidade de ajuste dos modelos se comporta, baseando-se na estatística de desvio absoluto média (DAM). As tabelas 16 e 17 contém informações sobre a qualidade do ajustamento dos modelos, para cada sub-amostra e exercício de obtenção de preferências (EVA e TTO), nas pesquisas das capitais (tabela 16) e de Minas Gerais (tabela 17). O número de exclusões (N exclusões) apresentado nas tabelas se refere ao número de estados de saúde, para analisar o número de exclusões de indivíduos em cada sub-amostra, deve-se dividir o número apresentado por 7, já que cada indivíduo avaliou 7 estados de saúde.

Tabela 16. Qualidade do Ajuste com Exclusão de Indivíduos Inconsistentes em cada Sub-Amostra, TTO, EVA: Amostra das Capitais

Amostra Total	Dist.>8	Dist.>7	Dist.>6	Dist.>5	Dist.>4	Todos inconsistentes	
EVA							
DAM	0,0804	0,0813	0,0808	0,0744	0,0707	0,0653	0,0595
N > 0.05	124	122	122	124	129	120	103
N > 0.10	67	68	67	56	54	47	43
N Exclusões	0	390	679	1764	2170	2513	11410
TTO							
DAM	0.0427	0.0432	0.0435	0.0426	0.0443	0.0444	0.0521
N > 0.05	80	83	86	78	80	93	103
N > 0.10	17	18	17	19	19	20	27
N Exclusões	0	161	581	1631	2597	3206	11032

NOTA: N>0,05: número de estados de saúde com desvio absoluto médio acima de 5%, N>0,10: número de estados de saúde com DAM acima de 10%. N exclusões: estados de saúde excluídos. FONTE: elaboração do autor,

Tabela 17. Qualidade do Ajuste com Exclusão de Indivíduos Inconsistentes em cada Sub-Amostra, TTO, EVA: Amostra de Minas Gerais

Amostra Total	Dist.>8	Dist.>7	Dist.>6	Dist.>5	Dist.>4	Todos inconsistentes	
EVA							
DAM	0.0356	0.0358	0.0369	0.0366	0.0369	0.0386	0.055
N > 0.05	22	23	24	22	22	27	45
N > 0.10	5	6	6	4	5	6	15
N Exclusões	0	49	308	707	882	2408	6545
TTO							
DAM	0.0393	0.0383	0.0378	0.0358	0.0349	0.0418	0.0443
N > 0.05	26	22	24	23	17	30	35
N > 0.10	4	4	4	3	3	4	8
N Exclusões	0	133	385	812	945	2212	7105

NOTA: N>0,05: número de estados de saúde com desvio absoluto médio acima de 5%, N>0,10: número de estados de saúde com DAM acima de 10%. N exclusões: estados de saúde excluídos.

Na pesquisa realizada nas capitais, a exclusão seletiva de indivíduos no TTO leva a Desvios Absolutos Médios (DAM) mais baixos conforme eliminam-se aqueles que cometem inconsistências entre pares de estados de saúde com distâncias maiores do que 6. Este resultado representa melhoria sutil em relação à utilização da amostra total (DAM da amostra total é de 0,0427 enquanto que DAM com exclusão de indivíduos que cometem inconsistências acima de 6 é de 0,0426). A figura 3 apresenta estes resultados, mostrando que os efeitos são mínimos. Ao remover todos os inconsistentes, o DAM apresenta uma piora considerável, ficando inclusive acima de 5%. O número de exclusões realizado para distâncias maiores do que 6 foi de 233 (=1631/7) observações, com 78 estados com $N < 0,05$ e 19 estados com $N > 0,10$. Na Figura 3 pode-se ver que os maiores desvios em reação ao predito encontram-se nas caldas da distribuição: estados muito graves acabam recebendo valorações muito negativas e estados mais leves recebem valorações mais próximas a 1.

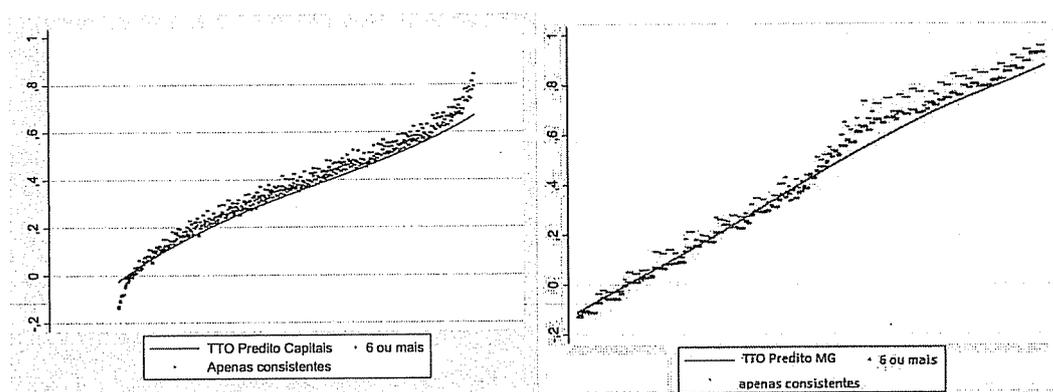


FIGURA 7. Comparação dos valores preditos para o TTO para a amostra total e para 2 sub-amostras- Pesquisa das capitais e do estado de MG

Na EVA, nenhuma das sub-amostras apresenta um desvio absoluto médio abaixo de 5%, refletindo uma baixa qualidade de ajuste destes modelos. A remoção seletiva entre os modelos sugere que o melhor ajuste possível é obtido com a exclusão de todos os inconsistentes, mas novamente não é um grau de ajuste aceitável. O número de indivíduos removidos chega a 1630 (=11410/7), sendo que 103 apresentam $n < 0,05$ e 27 apresentam $n > 0,10$.

Os resultados obtidos em Minas Gerais mostram a exclusão seletiva de indivíduos leva a Desvios Absolutos Médios (DAM) mais baixos para o TTO com a exclusão daqueles que

cometem inconsistências entre pares de estados de saúde com distâncias maiores do que 5. A melhoria em relação à utilização da amostra total ($DAM = 0,0393$) é maior do que aquela verificada na amostra das capitais (DAM considerando exclusão de inconsistências entre pares com distância maior que 5 é de $0,0358$). A figura 4 apresenta estes resultados, mostrando estes efeitos. Ao remover todos os inconsistentes, o DAM também apresenta uma piora considerável ($DAM = 0,043$). O número de exclusões realizado foi de 135 (=945/7) observações, com 17 estados com $N < 0,05$ e 3 estados com $N > 0,10$.

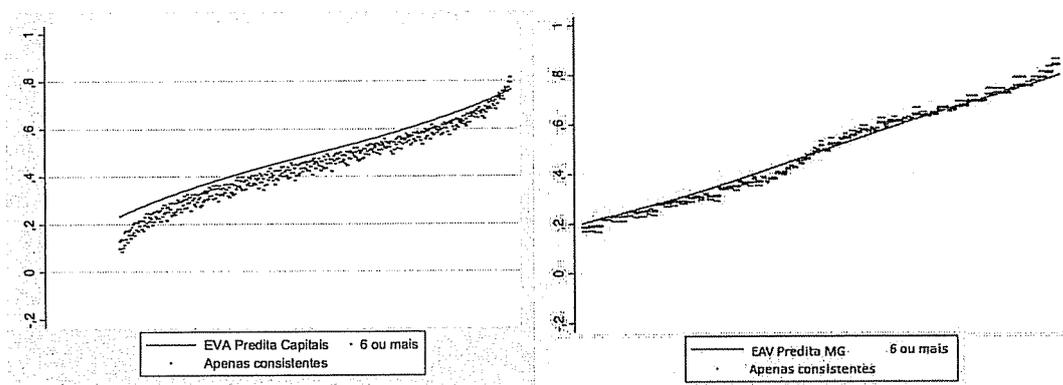


FIGURA 8. Comparação dos valores preditos para a EVA para a amostra total e para 2 sub-amostras- Pesquisa das capitais e Estado de Minas Gerais

Na EVA, não é indicada a remoção de observações, a maior qualidade de ajuste verifica-se com a amostra total ($DAM = 0,0356$), sendo que o pior ajuste se verifica quando todos os inconsistentes são excluídos ($DAM = 0,055$).

Estes resultados indicam que a utilização de blocos de estados fixos balanceados traz benefícios em termos de qualidade de ajuste dos modelos de regressão para a estimação dos pesos sociais por estados de saúde. As comparações aqui apresentadas revelam que a opção por coletar as informações utilizando um conjunto menor de estados de saúde (102 estados) pode produzir resultados com melhor ajuste do que a pesquisa que coleta informações para os 243 estados de saúde oriundos do EQ-5D. Mesmo após a exclusão de respostas inconsistentes e a avaliação ponderada pelas distâncias, os resultados produzidos na amostra do Estado de Minas Gerais apresentam menores taxas de inconsistência e maior qualidade de ajuste para a estimação das preferências sociais por estados de saúde.

Este trabalho não exaure o tema das inconsistências lógicas, mas propõe uma avaliação em relação a sua natureza, e sobre quais as características estão associadas à sua ocorrência. Pode-se ver também que a escolha por um desenho de estudo com estados de saúde balanceados é capaz de produzir resultados adequados, sem a necessidade de coletar informações para todos os estados de saúde.

Uma possível agenda de pesquisa futura pode se debruçar sobre como a ordem de apresentação dos estados de saúde pode afetar a ocorrência dos desvios lógicos. A ocorrência de um possível efeito de enquadramento pode ser um indutor destes desvios, uma vez que conforme estados muito próximos são apresentados, a distinção entre eles pode se tornar difusa.

Além disso, outro ponto importante a ser investigado é se os entrevistadores podem afetar de alguma forma a percepção e a análise dos respondentes. Na pesquisa realizada no estado de Minas Gerais, Andrade *et al.* (2005), indicam o entrevistador pode afetar a quantidade de inconsistências cometidas. Os principais fatores devem estar relacionados à explicação dos exercícios e a condução da pesquisa.

6. Considerações Finais

Este trabalho tem como objetivo comparar duas pesquisas voltadas para a estimação das preferências sociais por estados de saúde no Brasil em uma oportunidade única com características populacionais muito próximas, mas com um desenho de estudo diferenciado. A comparação das taxas de inconsistências pôde ser realizada nos três exercícios previstos no protocolo de aplicação do EQ-5D, tornando possível a investigação sobre o potencial efeito da diferente metodologia de seleção dos estados de saúde apresentados aos indivíduos. A análise também foi aprofundada para investigar o comportamento das inconsistências lógicas em relação à distância entre os pares de estados de saúde.

O exercício de valoração de estados de saúde é uma atividade complexa e cognitivamente desafiadora, demandando diferentes habilidades de comparação, reflexão e avaliação por parte dos entrevistados. Verificou-se que a dificuldade encontrada na avaliação depende não somente do exercício de valoração proposto, mas também da descrição dos estados de saúde. Neste trabalho as inconsistências lógicas também foram analisadas em termos de distâncias entre pares comparáveis e os resultados mostram que os desvios lógicos se concentram entre pares de estados próximos. Apesar de existir uma ordenação lógica intrínseca entre os estados de saúde, perceber diferenças sutis na descrição dos estados de saúde não é uma atividade trivial.

A presença de avaliações inconsistentes em estudos de estimação de preferências sociais por estados de saúde é recorrente não existe ainda um consenso na literatura a respeito de seu tratamento. Os efeitos encontrados neste trabalho mostram que a maioria dos desvios lógicos ocorre entre estados de saúde próximos, o que reflete que os indivíduos aparentemente compreenderam a atividade de avaliação adequadamente, mas que cometeram desvios lógicos ao não identificar a diferença da descrição dos estados de saúde.

O percentual de entrevistados que comete pelo menos um desvio lógico é de aproximadamente 50%, nas três formas de obtenção das preferências por estados de saúde e nas duas pesquisas. Já as taxas de inconsistências calculadas foram mais elevadas do que outros estudos realizados na Espanha (BADIA, HERDMANN e MONSERRAT, 1999), Dinamarca (WITTRUP-JENSEN; PEDERSEN e LAURIDSEN, 2008) e Holanda (KRABE, ESINK-BOT e BONSEL, 1997). Deve-se destacar que a heterogeneidade em termos

socioeconômicos presente no Brasil e o alcance destas duas pesquisas são maiores do que os estudos realizados nestes países europeus. Além disso, o desenho do estudo implementado no Brasil apresenta 9 estados de saúde para avaliação individual, organizados em blocos fixos e balanceados no caso de Minas Gerais e de forma aleatória, visando coletar informações para os 243 estados de saúde na pesquisa realizada nas capitais.

Os estudos europeus acima citados coletam informações para, pelo menos, 13 estados de saúde, o que pode afetar as taxas de inconsistência lógica através de dois mecanismos com efeitos distintos: o processo de aprendizagem interno aos exercícios e o possível esgotamento dos entrevistados ao realizar a pesquisa.

O aprendizado interno ao exercício é algo desejável, que pode diminuir o número de avaliações inconsistentes conforme o indivíduo se familiariza com a descrição dos estados de saúde e pode perceber com maior clareza a diferença entre cada condição proposta. Na via oposta, quando são apresentadas quantidades elevadas de estados de saúde, a realização de três exercícios de valoração seguida pela coleta de informações pessoais pode tornar a entrevista muito extensa e prejudicar a coleta de dados. O enfado ou exaustão provocados no entrevistado podem incentivar comportamentos que prejudicam a aferição das preferências, conforme Craig e Ramachandran (2006) definem como “sabotadores”.

Em termos das taxas de inconsistências médias ponderadas, os exercícios do ranking e do *Time Trade Off* apresentaram valores mais elevados do que a Escala Visual Analógica, em ambas as pesquisas. O primeiro exercício previsto no protocolo de aplicação MVH é o ordenamento (*ranking*), visando familiarizar os indivíduos com a descrição dos estados de saúde e permitir uma análise comparativa entre eles. As taxas encontradas para este exercício são mais elevadas do que as taxas da escala visual analógica, já que o ordenamento é a primeira etapa da entrevista.

Nas duas pesquisas realizadas no Brasil foi implementada uma alteração metodológica ao protocolo MVH original: os cartões que representam os estados de saúde foram recolhidos e embaralhados entre a execução do ranking e da escala visual analógica. Esta alteração produz diferenças das taxas de inconsistência entre os dois exercícios, uma vez que diminui a correlação entre as avaliações (que deveria ser próxima de 1) e permite captar melhor as singularidades de cada exercício.

As maiores taxas de inconsistências verificadas no TTO neste trabalho, em ambas as pesquisas decorrem da maior complexidade do exercício, além do fato de que cada estado de saúde ser comparado apenas com os estados âncora (saúde perfeita ou morte). Como cada avaliação acontece individualmente em relação aos estados âncora é possível que o indivíduo não se recorde plenamente das avaliações anteriores, o que pode aumentar o número de estados avaliados de forma contrária à sua estrutura lógica.

Os estados de saúde são apresentados de maneira sequencial, o que também pode afetar sua avaliação. Caso o indivíduo se depare com estados de saúde muito próximos entre si, pode acabar atribuindo valores iguais para condições de saúde logicamente inferiores. As maiores taxas de inconsistência encontradas na amostra das capitais pode ser um indício deste fenômeno, já que a escolha aleatória dos estados de saúde pode produzir um conjunto de estados com descrição parecida. Este fenômeno não ocorre na pesquisa realizada no estado de Minas Gerais, que sempre apresenta 2 estados graves, 2 estados moderados e 2 estados leves aos indivíduos, visando deixar uma distinção mais aparente entre as condições de saúde.

No protocolo de aplicação utilizado em ambas as pesquisas os estados de saúde eram apresentados aos indivíduos em nove cartões. Cada um continha a descrição das dimensões analisadas no EQ-5: mobilidade, atividades usuais, cuidados pessoais, ansiedade ou depressão e dor ou desconforto. A estrutura de todos os cartões mantinha esta mesma sequência de dimensões, para que a comparação entre elas ficasse mais evidente aos entrevistados. É possível que esta distribuição possa induzi-los a comparar mais criticamente as dimensões de saúde que aparecem primeiro (mobilidade, atividades usuais e cuidados pessoais). Entretanto, essa discussão tem um caráter mais profundo, uma vez que se tratam de preferências individuais. É natural que algumas dimensões sejam mais relevantes para determinados grupos sociais, já que os indivíduos enfatizam as dimensões que lhes são mais importantes. Para populações rurais, restrições na mobilidade são mais impactantes em termos de qualidade de vida do que para populações urbanas, que podem dispor de melhores condições de acessibilidade.

Uma proposta para avaliar como os indivíduos reagem em termos de inconsistências lógicas seria alterar a estrutura descritiva dos cartões que contém os estados de saúde. Alterar a ordem das dimensões de saúde pode forçar os indivíduos a analisar mais profundamente cada cartão.

Esta investigação não é possível neste trabalho, mas pode ser aprofundada em pesquisas posteriores, uma vez que demanda um desenho mais complexo, que permita a alteração da ordem como as dimensões de saúde estão organizadas e também a ordem de apresentação do conjunto de estados de saúde.

Como visto, a complexidade da avaliação dos estados de saúde decorre tanto dos exercícios de avaliação, como também de suas próprias características. A avaliação das inconsistências proposta neste trabalho também considerou a distância entre os pares de estados de saúde. Os resultados apontam, em ambas as pesquisas e nos três exercícios de obtenção de preferências, que a ocorrência das mesmas decai conforme a distância entre os estados de saúde se eleva. Este é um resultado esperado e desejável e indica que ao menos parte dos desvios lógicos ocorre pela dificuldade dos respondentes em diferenciar estados com diferenças em poucas dimensões de saúde.

As diferenças encontradas entre a pesquisa das capitais e do Estado de Minas Gerais indicam que as inconsistências se concentram em pares de estados de saúde mais próximos. Em termos de taxas de inconsistências, o que se verifica é que o estudo de Minas Gerais produz taxas mais baixas, considerando todas as distâncias, nos três exercícios apresentados. Este resultado é importante e revela que estrutura de blocos de estados fixa balanceada se apresenta como uma alternativa viável para a estimação das preferências sociais por estados de saúde, sem a necessidade de coletar informação para a totalidade dos 243 estados.

A estimação dos pesos sociais por estados de saúde produzida nas duas pesquisas também aponta que os resultados produzidos no estado de Minas Gerais têm melhor grau de ajuste estatístico tanto na escala visual analógica quanto no *Time Trade Off*. Este trabalho também propõe a exclusão seletiva de entrevistados inconsistentes baseado na distância em que os desvios lógicos foram cometidos. Foram excluídos primeiramente aqueles indivíduos que cometeram valorações inconsistentes entre estados de saúde mais distantes, o que é um possível indicativo que não compreenderam a atividade de avaliação.

A exclusão seletiva de indivíduos inconsistentes produziu estimativas com melhor grau de ajuste no caso da pesquisa de Minas Gerais, tanto para o TTO quanto para a escala visual analógica. A remoção de indivíduos que comete inconsistências entre estados de saúde com

distâncias iguais ou superiores a 6 melhorou significativamente a qualidade dos modelos estatísticos.

No caso das capitais, a exclusão seletiva de indivíduos inconsistentes não alterou de forma significativa o grau de ajuste estatístico dos modelos para o TTO. Para a escala visual analógica, o melhor grau de ajuste é produzido quando são excluídos todos os indivíduos inconsistentes.

Existe um debate na literatura a respeito do tratamento de respondentes inconsistentes, enquanto Craig e Ramachandran (2006) e Ohinmaa e Sintonen (1999) defendem a exclusão destes indivíduos, Dolan (1997) sugere que pode existir um possível trade-off entre representatividade e qualidade das preferências obtidas na pesquisa. Este trabalho propõe a exclusão baseada em critérios mais refinados, de indivíduos que possivelmente não entenderam adequadamente os exercícios de valoração.

O que se verifica com a comparação das duas pesquisas é que a utilização de blocos fixos de estados de saúde produziu proporções de respondentes inconsistentes similares, mas com taxas de inconsistências inferiores e maior qualidade de ajuste para modelos de estimação dos pesos sociais por estados de saúde. Deve-se levar em conta também que uma pesquisa que busca estimar pesos utilizando a totalidade de estados oriundos do EQ-5D possui uma logística mais custosa sem produzir necessariamente resultados mais satisfatórios. Para estudos futuros, sugerimos a utilização da estrutura de blocos fixos de estados de saúde, com uma cobertura mais ampla, visando obter maior representatividade possível para a obtenção das preferências sociais por estados de 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.

ANDRADE, M. *et al.* Logical Inconsistencies in 3 Preference Elicitation Methods for EQ-5D Health States: A Study in the Brazilian Population. *Medical Decision Making*, v. 36, p. 242-252, Feb. 2016.

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.

AUGUSTOVSKI, F. *et al.* Argentine Valuation of the EQ-5D Health States. *Value In Health*, v. 12, n. 4, p. 587-596, 2009.

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.

BAGUST, A. Improving valuation sampling of EQ-5D health states. *Health and Quality of Life Outcomes*, v. 11, n. 14, 2013.

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. 2012a.

BANTA, D.; LUCE, B. Assessing the cost-effectiveness of prevention. *Journal of Community Health* v. 9, n. 2, 1983.

BATTISTA, R.; HODGE, M.. The evolving paradigm of health technology assessment: reflections for the millennium. *Canadian Medical Association Journal*, v. 110, n. 20, p.1464-1467.

BRASIL. Secretaria de Políticas de Saúde. Política Nacional de Medicamentos. *Revista de Saúde Pública*, São Paulo, v. 34, n. 2, p. 206-209. Apr. 2000.

BRASIL. MINISTÉRIO DA SAÚDE. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. Avaliação de tecnologias em saúde: institucionalização das ações no Ministério da Saúde. *Revista de Saúde Pública*, Brasília, v. 40, n. 4, p. 743-747, 2006.

BRASIL. MINISTÉRIO DA SAÚDE. Secretaria-Executiva. Área de Economia da Saúde e Desenvolvimento. *Avaliação econômica em saúde: desafios para gestão no Sistema Único de Saúde*. Brasília: Editora do Ministério da Saúde, 2008. 104 p.

BRASIL. MINISTÉRIO DA SAÚDE. Secretaria-Executiva. Área de Economia da Saúde e Desenvolvimento. *Avaliação de tecnologias em saúde: ferramentas para a gestão do SUS*. Brasília: Editora do Ministério da Saúde, 2009a. 110 p.

- BRAZIER, J. *et al.* Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal*, v. 305, p. 160-164, Jul. 1992.
- BRAZIER, J. *et al.* A review of the use of health status measures in economic evaluation. *Health Technology Assessment*, v. 3, n. 9, p. 1-163, 1999.
- BRAZIER, J.; ROBERTS, J.; DEVERILL, M. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, v. 21, p. 271-292, 2002.
- BRAZIER, J. *et al.* Comparison of health states utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. *Eur. J. Health Econ.*, v. 13, n. 5, p. 575-587, Oct. 2012.
- BUSSCHBACH, J. *et al.* Observations on one hundred students filling the EuroQol questionnaire. In: KIND, P.; *et al.* *EQ-5D Concepts and Methods*, p. 81-90, 2005.
- CAMARGOS, M.; PERPÉTUO, I.; MACHADO, C. Expectativa de vida com incapacidade funcional em idosos em São Paulo, Brasil. *Rev. Panam. Salud Publica*. v. 17, n. 5/6, p. 379-86. 2005.
- CAMERON, A.; TRIVEDI, P. *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press, 2005. 1034 p.
- CAMPOLINA, A.; CICONELLI, R. Qualidade de vida e medidas de utilidade: parâmetros clínicos para as tomadas de decisão em saúde. *Rev Panam Salud Publica*, v. 19, n. 2, p. 128-136, 2006.
- CARVALHO, G. I.; SANTOS, L. Sistema Único de Saúde: comentários à Lei Orgânica da Saúde (Leis 8.080/90 e 8.142/90). 2.ed. *São Paulo*: Hucitec, 1995.
- CHEVALIER, J.; POUVOURVILLE, G. Valuing EQ-5D using Time Trade-Off in France. *The European Journal of Health Economics*, V. 14, P. 57-66, Sep. 2011.
- 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.
- COAST, J. *et al.* Investigating Choice Experiments for Preferences of Older People (ICEPOP): evaluative spaces in health economics. *Journal of Health Services Research & Policy*, v. 13, n. 3, p. 31-37, Oct. 2008.
- COSTA, R.; BARBOSA, R.; ZUCCHI, P. Expenditures in the health care system in Brazil: the participation of states and the Federal District in financing the health care system from 2002 to 2013. *CLINICS*, v. 70, n. 4, p. 237-241. 2015.
- 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.

CRAIG, B. The Duration Effect: A Link Between TTO And VAS Values. *Health Economics*, v. 18, p. 217-225, 2009.

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; BUSSCHABACH, J; SALOMON, J. Modeling Ranking, Time Trade-Off, and Visual Analog Scale Values for EQ-5D Health States: A Review and Comparison of Methods. *Medical Care*, v. 47, n. 6, p. 634-641, Jun. 2009b.

CRAIG, B. *et al.* Modeling Ranking, Time Trade-off and Visual Analogue Scale Values for EQ-5D Health States: a review and comparison of methods. *Medical Care*, v. 47, n. 6, p. 634-641, Jun. 2009.

CRUZ, L. *Medidas de Qualidade de Vida e Utilidade em uma Amostra da População de Porto Alegre*. 2010. 270 f. Tese (Doutorado em Epidemiologia) – Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010.

CRUZ, L. *et al.* Estimating the SF-6D Value Set for a Population-Based Sample of Brazilians. *Value in health*, v. 14, n. Suppl 1, p. S108-S114, Jul-Aug. 2011.

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.

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.

DOLAN, P.; GUDEX, C. Time preference, duration and health state valuations. *Health Economics*, v. 4, n. 4, p. 289-299, Jul-Aug. 1995.

DOLAN, P.; KIND, P. Inconsistency and Health State Valuations. *Soc. Sci. Med.*, v. 42, n. 4, p. 609-615, 1996.

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. *et al.* Valuing health states: A comparison of methods. *Journal of Health Economics*, vol. 15, p. 209-231, 1996.

DOLAN, P. Modelling valuations for health states: the effect of duration. *Health Policy*, v. 38, p. 189-203, 1996.

DOLAN, P. The Measurement of Health-related Quality of Life for use in Resource Allocation Decision in Health care. In: CULYER, A.; NEWHOUSE, J (Ed.). *Handbook of Health Economics*, Elsevier, 2000. v. 1B, chapter. 32, p. 1723-1760.

DOLAN P.; ROBERTS, J. To what extent can we explain time trade-off values from other information about respondents? *Social Science & Medicine*, v. 54, p. 919-929. 2002.

DRUMMOND, M.; STODDART, G.; TORRANCE, G. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Medical Publications, 2nd ed., 1997.

ESSINK-BOT, M.; STOUTHARD, M.; BONSEL, G. Generalizability of valuation on health states collected with the EuroQol-questionnaire. *Health Economics*, v.2, p. 237-246, 1993.

FEENY, D. *et al.* Multiattribute and Single-Attribute Utility Functions for the Health Utilities Index Mark 3 System. *Medical Care*, v. 40, n. 2, p. 113-128, 2002.

GOLD, M.; FRANKS, P.; ERICKSON, P. Assessing the Health of a Nation: The predictive validity of a preference-based measure and self-related health. *Medical Care*, vol. 34, n. 2, p. 163-177. Feb. 1996.

GOLICKI, D.; *et al.* Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQVAS in stroke patients. *Quality of Life Research*, v. 25, p. 1555-1563, 2015.

GREINER, W.; *et al.* Validating the EQ-5D with time trade off for the German population. *European Journal of Health Economics*, v. 6, p. 124-130, 2005.

GUDEX, C. *Time Trade-Off User Manual: Props and Self-Completion Method*. Centre for Health Economics, University of York, Occasional Paper Series, 1994.

HAILEY, D.M. Health technology assessment in Canada: diversity and evolution. *The Medical Journal of Australia*, v. 187, n. 5, p. 286-88. Sep. 2007.

HAUSMAN, J.; McFADDEN, D. Specification Tests for the Multinomial Logit Model. *Econometrica*, v. 52, n. 5, p. 1219-1240, Sep. 1984.

HERDMAN, M.; *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, v. 20, p. 1727-1736. 2011.

HURLEY, J. An Overview of the normative economics of the health sector. In: CULYER, A.; NEWHOUSE, J. *Handbook of Health Economics*. Elsevier, 2000. v. 1, chapter 2, p. 56-118.

HURST, N.; RUTA, D.; KIND, P. Comparison of the Mos Short Form-12 (SF12) health status questionnaire with the SF36 in patients with rheumatoid arthritis. *British Journal of Rheumatology*, v. 37, n. 8, p. 862-869, 1998.

JACKSON, T. Health technology assessment in Australia: challenges ahead. *The Medical Journal of Australia*, v. 187, n. 5, p. 263-64. Sep. 2007.

KAHNEMAN, D.; TVERSKY, A. Prospect theory: an analysis of decision under risk. *Econometrica*, v. 47, p. 263-291, 1979.

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.

KLARMAN, H.; FRANCIS, J.; ROSENTHAL, G. Cost-effective analysis applied to the treatment of chronic renal disease. *Medical Care*, v. 6, n. 1, p. 46-54. Jan-Feb. 1968.

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.

KRABBE, P. *et al.* Testing the interval-level measurement property of multi-item visual analogue scales. *Quality of Life Research*, v. 15, n. 10, p. 1651-1661, Dec. 2006.

KRAUS-SILVA, L. Avaliação tecnológica e análise custo-efetividade em saúde: a incorporação de tecnologias e a produção de diretrizes clínicas para o SUS. *Ciência e Saúde Coletiva*, v. 8, n. 2, p. 501-520, 2003

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.

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.

LEE, Y. *et al.* South Korean time trade-off values for EQ-5D health states: modeling with observed values for 101 health states. *Value in Health*, v. 12, n. 8, p. 1187-1193, 2009.

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.

LOOMES, G.; MCKENZIE, L. The Use of QALYs in Health Care Decision Making. *Soc. Sci. Med.*, v. 28, n. 4, p. 299-308, 1989.

MACRAN, S.; KIND, P. Valuing EQ-5D health states using a modified MVH protocol: preliminary results. *EuroQol Group 16th Plenary Meeting, Spain*, 2000.

MACRAN, S.; WEATHRLY, H.; KIND, P. Measuring Population Health: A Comparison of Three Generic Health Status Measures. *Medical Care*, v. 41, n. 2, p. 218-231, 2003.

Médici, A. Economia e financiamento do setor saúde no Brasil: balanços e perspectivas do processo de descentralização. *Faculdade de Saúde Pública da USP*, São Paulo. 1994.

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE - NICE. Guide to the methods of technology appraisal. 2008 Jun. Available from: <<http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf>>.

NEUMANN, P. *Using cost-effectiveness analysis to improve health care: opportunities and barriers*. Oxford University Press US. 2005.

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

NORMAN, R.; *et al.* International Comparisons in Valuing EQ-5D Health States: A Review and Analysis. *Value in Health*, v. 12, n. 8, p. 1194-1200, 2009.

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

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO - PNUD. Atlas do Desenvolvimento Humano no Brasil 2003. Disponível em: <<http://www.pnud.org.br/atlas>>.

OHINMAA, A.; SINTONEN, H. Inconsistencies and modelling the Finnish Euroqol (EQ-5D) Preference Values. EuroQol meetings. 1999.

OHINMAA, A.; *et al.* Comparison of Finnish and U.S.-based Visual Analog Scale Valuations of the EQ-5D Measure. *Medical Decision Making*, v. 20, p. 281-289, 2000.

ROBINSON, A.; LOOMES, G.; JONES-LEE, M. Visual Analog Scales, Standard Gambles, and Relative Risk Aversion. *Medical Decision Making*, v. 21, n. 1, p. 17-27, Jan-Feb. 2001.

ROMERO, D.; LEITE, I.; SZWARCOWALD, C. Healthy life expectancy in Brazil: applying the Sullivan method. *Cadernos de Saúde Pública*, Rio de Janeiro, v. 21, Sup:S7-S18, 2005.

SANTOS, M. *et al.* Brazilian Valuation of EQ-5D-3L Health States: Results from a Saturation Study. *Medical Decision Making* 0272989X15613521, publicado em 22/10/2015.

SACKETT, D.; TORRANCE, G. The utility of different health states as perceived by the general public. *Journal of Chronic Disease*, v. 31, n. 11, p. 697-704, 1978.

SECOLI, S. *et al.* Farmacoeconomia: perspectiva emergente no processo de tomada de decisão. *Ciências e Saúde Coletiva*, v. 10, n. Sup, p. 287-296. Set-Dez. 2005.

SHAW, J.; COONS, S.; JOHNSON, J. US Valuation of the EQ-5D Health States: Development And Testing Of The D1 Valuation Model. *Medical Care*, v. 43, n. 3, p. 203-220, 2005.

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

TORRANCE, G.; THOMAS, W.; SACKETT, D. A utility maximization model for evaluation of health care programs. *Health Services Research*, v. 7, n. 2, p. 118-33, 1972.

TORRANCE, G. Measurement of Health State Utilities for Economic Appraisal: a Review. *Journal of Health Economics*, v. 5, p. 1-30, 1986.

TUNIS, S.; KANG, J. Improvements in medicare coverage of new technology: how medicare has responded to the need to improve access to beneficial technologies. *Health Affairs*, v. 20, n. 5, p. 83-85, 2001.

von NEUMANN, J.; MORGENSTERN, O. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

WARNER, K. E.; LUCE, B. R. Cost-benefit and cost-effectiveness analysis in health care. *Ann Arbor: Health Administration Press*, 1982.

WILLE, N.; *et al.* Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research*, v. 19, p. 875- 886. 2010.

WILLIAMS, A. "The Value of QALYs". *Health and Social Service Journal (supplement)*, 1985

WITTRUP-JENSEN, K.; PEDERSEN, J.; LAURIDSEN, K. Modelling Danish EuroQol (EQ-5D) Tariffs by Applying the Time Trade-Off Method. *Health Economic Papers, University of South Denmark*, v. 4, 2008.

WORLD HEALTH ORGANIZATION. Global Health and Aging. Disponível em: <http://www.who.int/ageing/publications/global_health/en/> Acesso em: 20/08/2015.

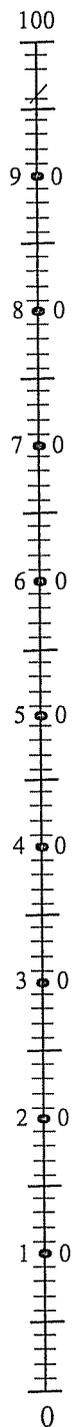
WORLD HEALTH ORGANIZATION. World Health Statistics 2014. Disponível em: <http://www.who.int/gho/publications/world_health_statistics/2014/en/> Acesso em: 23/09/2015.

ZARATE, V. *et al.* Social valuation of EQ-5D health states: the Chilean case. *Value In Health*, v. 14, p. 1135-41, 2011.

ANEXO A - INSTRUMENTOS

O melhor estado de
saúde imaginável

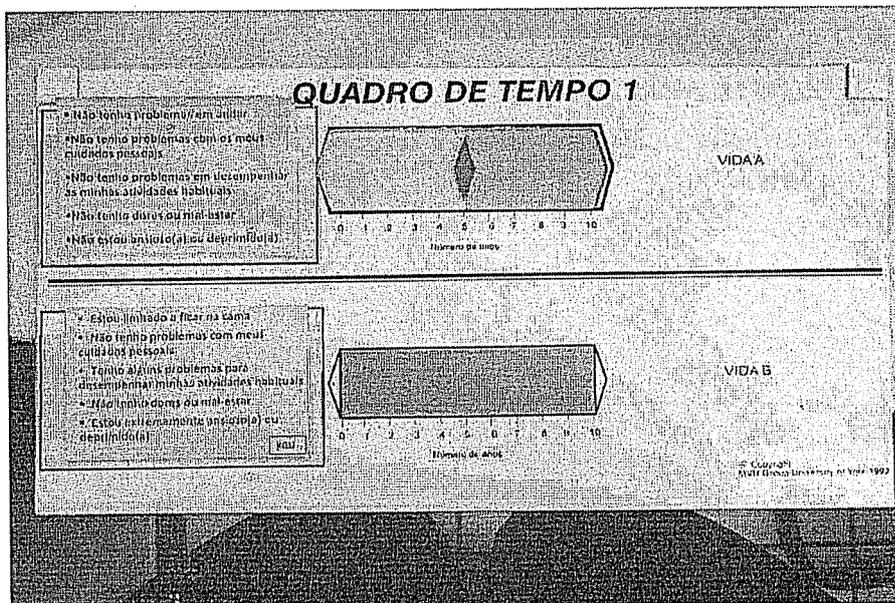
Escola Analógica Visual



O pior estado de
saúde imaginável

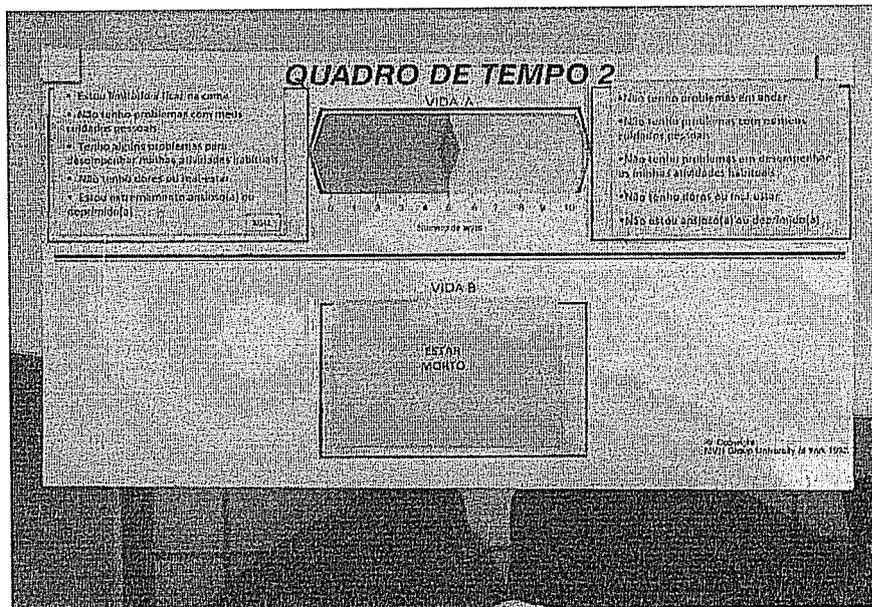
ANEXOB – QUADRO DA TROCA DE TEMPO

FIGURA A – QUADRO DE TEMPO 1: Estados melhores que a morte



Fonte: Mesquita, 2013.

FIGURA B – QUADRO DE TEMPO 2: Estados piores que a morte



Fonte: Mesquita, 2013.

SISTEMA DESCRITIVO EQ-5D

Assinale com uma cruz (assim ☒), um quadrado de cada um dos seguintes grupos, indicando qual das afirmações melhor descreve o seu estado de saúde **atual**.

Mobilidade

- Não tenho problemas em andar
- Tenho alguns problemas em andar
- Estou limitado a ficar na cama

Cuidados Pessoais

- Não tenho problemas com os meus cuidados pessoais
- Tenho alguns problemas para me lavar ou me vestir
- Sou incapaz de me lavar ou vestir sozinho

Atividades Habituais (*ex. trabalho, estudos, atividades, domésticas, atividades em família ou de lazer*)

- Não tenho problemas em desempenhar as minhas atividades habituais
- Tenho alguns problemas em desempenhar as minhas atividades habituais
- Sou incapaz de desempenhar as minhas atividades habituais

Dor/Mal Estar

- Não tenho dores ou mal-estar
- Tenho dores ou mal-estar moderados
- Tenho dores ou mal-estar extremos

Ansiedade/Depressão

- Não estou ansioso(a) ou deprimido(a)
- Estou moderadamente ansioso(a) ou deprimido(a)
- Estou extremamente ansioso(a) ou deprimido(a)

SISTEMA DESCRITIVO EQ-5D

- Eu vou fazer algumas perguntas para você sobre diferentes estados de saúde e doença.
- Não existem respostas certas ou erradas, eu apenas gostaria de saber o que você pensa.
- Tudo o que você disser nesta entrevista será mantido em sigilo.

ENTREGUE O SISTEMA DESCRITIVO EQ-5DAO ENTREVISTADO E PEÇA AO ENTREVISTADO QUE O PREENCHA

1.

- Primeiramente, eu vou lhe pedir para que pense sobre seu estado de saúde atual.
- Marque um X em cada um dos grupos para mostrar quais situações melhor descrevem sua saúde atual.
- Assinale com uma cruz (assim) , um quadrado de cada um dos seguintes grupos, indicando qual das afirmações melhor descreve o seu estado de saúde **atual**.

PEGUE O ENVELOPE DO GRUPO DE CARTÕES E EMBARALHE-OS.

ORDENAÇÃO

2.

- Aqui estão alguns cartões, cada um descreve um estado de saúde.
- Cada estado de saúde é descrito por uma combinação de cinco características deste questionário que você utilizou para avaliar o seu estado de saúde e o cartão referente a “ESTAR MORTO”.
- Os cartões não estão ordenados.
- Eu quero que você os ordene, deixando na parte superior o estado de saúde que você avalia ser o melhor, e na parte inferior o estado de saúde que você avalia como sendo o pior.
- Se você acha que alguns destes estados são iguais, posicione-os um ao lado do outro.
- Para cada um desses estados, eu quero que você se imagine nestas situações.
- Para cada estado, pense que a situação descrita vai durar por 10 anos, sem qualquer mudança, e depois você morre.
- Eu gostaria que você os ordenasse juntamente com o cartão em que está escrito “ESTAR MORTO”.
- Se você achar que algum estado é pior do que estar morto, posicione este cartão abaixo do cartão “ESTAR MORTO”.
- Agora, por favor, ordene os cartões.

QUANDO O ENTREVISTADO TERMINAR DE ORDENAR OS CARTÕES, VERIFIQUE SE ELE ENCONTRA-SE SATISFEITO COM A ORDENAÇÃO.

- Você está satisfeito com a sua ordenação? Se não estiver, fique à vontade para alterá-la.

DEIXE OS CARTÕES SOBRE A MESA EM SUA ORDEM DE CLASSIFICAÇÃO.

REGISTRE A ORDEM DOS ESTADOS NA COLUNA COM CABEÇALHO “MELHOR SITUAÇÃO”, NO QUADRO DE CLASSIFICAÇÃO I.

A) SE SOMENTE UM CARTÃO É ALOCADO PARA DETERMINADA POSIÇÃO, ESCREVA O CÓDIGO DE LETRAS CORRESPONDENTE NO ESPAÇO TRIPLO.

B) SE DOIS OU MAIS CARTÕES SÃO ALOCADOS NA MESMA POSIÇÃO, ESCREVA OS CÓDIGOS DE LETRAS DESSES CARTÕES NA MESMA LINHA.

QUADRO DE CLASSIFICAÇÃO I		
MELHOR SITUAÇÃO	MELHOR SITUAÇÃO	MELHOR SITUAÇÃO
1. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
2. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
3. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
4. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
5. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
6. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
7. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
8. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
9. <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
PIOR SITUAÇÃO	PIOR SITUAÇÃO	PIOR SITUAÇÃO

RECOLHA OS CARTÕES DA MESA.

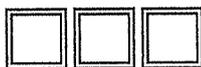
ESCALA ANALÓGICA VISUAL (EAV)

ENTREGUE A ESCALA ANALÓGICA VISUAL AO ENTREVISTADO

3.

- Eu gostaria que você olhasse esta escala.
- Como você pode ver, a extremidade superior marca 100.
- Este é o melhor estado de saúde possível que você pode imaginar.
- A extremidade inferior marca 0 e este é o pior estado de saúde possível que você pode imaginar.
- Agora eu gostaria que você marcasse na escala o quão bom ou ruim você pensa ser seu estado de saúde atual.
- Escreva "EU" próximo a este traço.

REGISTRE A PONTUAÇÃO QUE O ENTREVISTADO MARCOU PARA AVALIAR SEU PRÓPRIO ESTADO DE SAÚDE.



RETIRE O CARTÃO "ESTAR MORTO", EMBARALHE OS DEMAIS CARTÕES E ENTREGUE-OS AO ENTREVISTADO

4.

- Agora você vai marcar onde estes estados deveriam estar na escala.
- Imagine ainda que você se encontra em cada um desses estados e que ele vai durar por 10 anos, sem qualquer mudança, e então você morre.
- Use a escala de modo que os estados que você considera parecidos tenham pontuações próximas e os estados que você considera muito diferentes entre si tenham pontuações mais distantes.
- Você pode dar a mesma pontuação para mais de um estado.
- Escreva o código de letras de cada cartão próximo a cada traço que você realizar.

DEPOIS QUE O ENTREVISTADO TERMINAR DE VALORAR OS CARTÕES ENTREGUE O CARTÃO "ESTAR MORTO".

5. Aqui está o cartão "ESTAR MORTO". Onde você o colocaria nesta escala? Você pode alterar o local na escala em que você colocou os outros estados, caso queira.

RECOLHA A ESCALA ANALÓGICA VISUAL

REGISTRE A PONTUAÇÃO DE CADA CARTÃO AO LADO DOS SEUS RESPECTIVOS CÓDIGOS DE LETRAS NA COLUNA EAV DO QUADRO DE CLASSIFICAÇÃO II.

QUADRO DE CLASSIFICAÇÃO II

REGISTRO DO CÓDIGO DO CARTÃO E PONTUAÇÃO NO ESPAÇO EAV.

	CÓDIGO DO CARTÃO	EAV
1.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
2.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
3.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
4.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
5.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
6.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
7.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
8.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
9.	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>

PROTOCOLO DA TROCA DE TEMPO

RETIRE OS CARTOES DE SAÚDE PERFEITA E ESTAR MORTO E EMBARALHE OS DEMAIS.

MOSTRE O LADO 1 DO QUADRO DA TROCA DE TEMPO

- Eu vou mostrar para você algumas descrições dos estados de saúde que você viu anteriormente.
- Eu quero que você imagine que VOCÊ está nesses estados, os quais podem permanecer por diferentes durações.
- Não há respostas certas ou erradas. Eu só vou querer saber qual VOCÊ escolhe.

MOSTRE O CARTÃO COM O ESTADO PARA O RESPONDENTE.

- Por favor, leia a descrição deste cartão.

REGISTRE O CÓDIGO DO CARTÃO REFERENTE A ESTE ESTADO DE SAÚDE NOS QUADRADOS ABAIXO

1)

**USE O QUADRO DA TROCA DE TEMPO “1”
INSIRA O CARTÃO NO BOLSO CORRESPONDENTE À “VIDA B” E MOVA O MARCADOR DO QUADRO DA TROCA DE TEMPO DA “VIDA A” PARA “0” ANO**

- Vida B significa viver 10 anos neste estado e depois você morre
- Vida A não dá a você nenhum tempo de vida – você morreria agora
- Qual das duas vidas você prefere? Vida A ou Vida B ou elas são a mesma coisa?

REGISTRE A RESPOSTA DO INDIVÍDUO COM UM X NOS CAMPOS AO LADO DO QUADRO ABAIXO

2)

Escolha	Significado	Ação	SEÇÃO	<input type="checkbox"/>
VIDA A	Estado Pior que a morte	Vá para Pior que a Morte	II	<input type="checkbox"/>
VIDA B	Estado melhor que a morte	Vá para Melhor que a Morte	I	<input type="checkbox"/>
Mesma Coisa	Ponto de Indiferença Estabelecido	FIM		<input type="checkbox"/>

**SE O INDIVÍDUO ESCOLHEU VIDA “A” VÁ PARA A SEÇÃO II DO QUESTIONÁRIO
SE O INDIVÍDUO ESCOLHEU VIDA “B” VÁ PARA A SEÇÃO I DO QUESTIONÁRIO
SE VIDA “A” E VIDA “B” REPRESENTAM A MESMA COISA PARA O INDIVÍDUO, ENCERRE A AVALIAÇÃO DESTE ESTADO E PASSE PARA O PRÓXIMO CARTÃO.**

QUADRO DE VALOR

VALOR (anos)	<input type="text"/>
--------------	----------------------

SEÇÃO I – ESTADO MELHOR QUE A MORTE

CONTINUE USANDO O QUADRO DE TROCA DE TEMPO 1

REGULE O MARCADOR DO QUADRO DA TROCA DE TEMPO PARA 5 ANOS ($t = 5$)

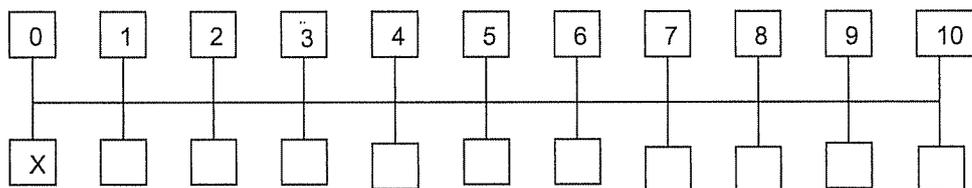
- Vida A significa que você vive por t anos em saúde perfeita e depois morre.
- Vida B significa que você vive por 10 anos nesse estado de saúde e depois morre.
- Você preferiria a Vida A ou Vida B ou elas são a mesma coisa?

Escolha	Significado	Ação
VIDA A	Prefere t anos em saúde perfeita a 10 anos em pior estado de saúde	Marque um \checkmark embaixo de t Mova o marcador do quadro da troca de tempo para DIMINUIR 1 ano ($\downarrow t$)
VIDA B	Prefere 10 anos em pobre estado de saúde a t anos em saúde completa	Marque um X embaixo de t Mova o marcador do quadro da troca de tempo para AUMENTAR 1 ano ($\uparrow t$)
A mesma coisa	Ponto de indiferença estabelecido	Marque um sinal de igualdade (=) embaixo de t

SE UM \checkmark E UM X OCORREREM UM AO LADO DO OUTRO ENCERRE A AVALIAÇÃO DESTE ESTADO DE SAÚDE E REGISTRE O VALOR DO PONTO MÉDIO NO QUADRO DE VALOR.

SE UM = FOR REGISTRADO ENCERRE A AVALIAÇÃO DESTE ESTADO DE SAÚDE E REGISTRE O VALOR DO t NO QUADRO DE VALOR.

CASO CONTRÁRIO REPITA COM O NOVO TEMPO t .



QUADRO DE VALOR

VALOR (anos)	
--------------	--

INICIE A AVALIAÇÃO DO PRÓXIMO ESTADO DE SAÚDE ATÉ COMPLETAR OS 7 ESTADOS DE SAÚDE

SEÇÃO II – ESTADO PIOR QUE A MORTE

PASSE A UTILIZAR O QUADRO DA TROCA DE TEMPO 2

RETIRE O CARTÃO DO ESTADO AVALIADO DO LADO 1 E PASSE-O PARA O LADO 2

REGULE O MARCADOR DO QUADRO DA TROCA DE TEMPO PARA 5 ANOS ($t = 5$)

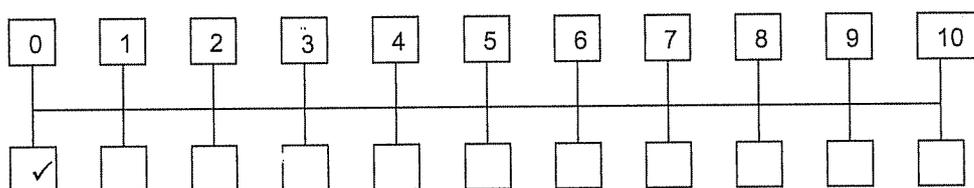
- Vida A significa que você viveria por t anos neste estado de saúde, seguido de $10-t$ anos em saúde perfeita.
- Vida B significa estar morto.
- Você preferiria a Vida A ou a Vida B ou elas são a mesma coisa?

Escolha	Significado	Ação
VIDA A	Prefere t anos em saúde ruim e acrescido de $10-t$ em saúde perfeita a estar morto	Marque um \checkmark embaixo de t Mova o marcador do quadro da troca de tempo para AUMENTAR 1 ano ($\uparrow t$)
VIDA B	Prefere estar morto a viver em estado de saúde ruim	Marque um X embaixo de t Mova o marcador do quadro da troca de tempo para DIMINUIR 1 ano ($\downarrow t$)
A mesma coisa	Ponto de indiferença estabelecido	Marque um sinal de igual (=) embaixo de t

SE UM \checkmark E UM X OCORREREM UM AO LADO DO OUTRO ENCERRE A AVALIAÇÃO DESTE ESTADO DE SAÚDE E REGISTRE O VALOR DO PONTO MÉDIO NO QUADRO DE VALOR.

SE UM = FOR REGISTRADO ENCERRE A AVALIAÇÃO DESTE ESTADO DE SAÚDE E REGISTRE O VALOR DO t NO QUADRO DE VALOR.

CASO CONTRÁRIO REPITA COM O NOVO TEMPO t .



QUADRO DE VALOR

VALOR (anos)	
--------------	--

INICIE A AVALIAÇÃO DO PRÓXIMO ESTADO DE SAÚDE ATÉ COMPLETAR OS 7 ESTADOS DE SAÚDE