



Ana Patrícia Ferreira Antunes

DECISÃO BASEADA EM PREFERÊNCIAS A MEDIÇÃO EM SAÚDE

Tese de Doutoramento em Gestão — Ciência Aplicada à Decisão,
orientada pelo Professor Doutor Pedro Lopes Ferreira e pela Professora Doutora Lara Noronha Ferreira
e apresentada à Faculdade de Economia da Universidade de Coimbra

Junho 2017



UNIVERSIDADE DE COIMBRA



• U • C •

FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Ana Patrícia Ferreira Antunes

Decisão baseada em preferências

A medição em saúde

Tese de Doutoramento em Gestão – Ciência Aplicada à Decisão,
apresentada à Faculdade de Economia da Universidade de Coimbra
para obtenção do grau de Doutor

Orientadores: Prof. Doutor Pedro Lopes Ferreira e Prof. Doutora Lara Noronha Ferreira

Coimbra, 2017

Ao Ricardo e à minha Mãe por estarem sempre ao meu lado e me apoiarem nas minhas escolhas.

AGRADECIMENTOS

Esta tese resulta do trabalho realizado nos últimos quatro anos e meio como aluna do Doutoramento em Gestão – Ciência Aplicada à Decisão, da Faculdade de Economia da Universidade de Coimbra. O documento agora apresentado é fruto do esforço individual mas seria impossível de concretizar sem o apoio das pessoas que me têm acompanhado ao longo da vida e de tantas outras que tive oportunidade de conhecer durante esta aventura.

Em primeiro lugar, gostaria de agradecer ao Professor Doutor Pedro Lopes Ferreira pela oportunidade que me concedeu em trabalhar neste projecto mas sobretudo pelo apoio e orientação que sempre me deu ao longo dos últimos 11 anos. Os seus conselhos permitiram-me encontrar um rumo profissional e os projectos de investigação conjunta têm sido uma fonte de aprendizagem e de inspiração. Hoje seria muito menos do que sou se não tivesse tido a oportunidade de o conhecer.

De seguida gostaria de agradecer à Professora Doutora Lara Noronha Ferreira pela experiência que partilhou e pelo que me ensinou.

Este projecto possibilitou conhecer uma vasta equipa internacional e, apesar de estar muito grata a todos, há três pessoas que gostaria de destacar: Arnd Jan Prause pela sua prontidão, apoio e paciência infinitos nas incontáveis ligações remotas e dúvidas técnicas sobre o EQ-VT, Kristina Ludwig pelo apoio técnico constante e Juan Manuel (Juanma) Ramos-Goñi por tudo o que me ensinou e pela motivação que transmitiu. É um privilégio poder ter contado com o apoio técnico desta superequipa do grupo EuroQol! Por outro lado, o projecto de valoração do EQ-5D-5L para a população portuguesa não seria possível sem o financiamento do grupo EuroQol e, por esse, motivo estou muito grata a Bernhard Slaap (EuroQol Executive Director) e Elly Stolk (EuroQol Scientific Team Leader) por esta oportunidade.

Gostaria de agradecer também ao Luís Guerra pelo trabalho fantástico que fez com a capa desta tese e pela paciência e ajuda na gestão deste curso de doutoramento com a minha componente profissional na Área Funcional dos Cuidados de Saúde Primários do Departamento de Planeamento e Contratualização da Administração Regional de Saúde do Centro. Agradeço também ao João Paulo Gonçalves por todo o apoio.

Queria ainda deixar o meu agradecimento à Filipa Landeiro pela amizade e apoio contínuos e pelas horas de conversa à distância.

Este trabalho não seria possível sem o contributo dos entrevistadores e dos participantes no estudo e, por isso, gostaria de prestar um agradecimento genuíno ao seu empenho e colaboração.

Por último, mas não menos importante, gostaria de agradecer à minha Mãe por sempre me ter apoiado e por acreditar nos meus sonhos e ao Ricardo por estar ao meu lado na jornada da vida.

Coimbra, Junho de 2017

FINANCIAMENTO

O estudo português de valoração do EQ-5D-5L, do qual decorre esta tese, foi financiado pelo grupo EuroQol.

We can only see a short distance ahead,
but we can see plenty there that needs to be done.
Alan Turing, *Computing machinery and intelligence* (1950)

RESUMO

O processo de decisão, num contexto de recursos limitados, tem consequências importantes na saúde dos indivíduos. A escolha de novas tecnologias ou de intervenções em saúde e a reavaliação de procedimentos já existentes são aspectos particularmente desafiantes para os decisores, sobretudo no contexto económico actual. O processo de decisão requer abordagens transparentes e racionais a fim de garantir a maximização da saúde e a redução das desigualdades. Os instrumentos de medição em saúde com recurso a métodos de elicitación de preferências são parte integrante das técnicas de avaliação económica. Estes instrumentos permitem o apoio ao processo de decisão, nomeadamente na análise custo-utilidade, através da obtenção de utilidades em saúde. A sua aplicação torna, também, possível ir ao encontro da perspectiva dos cidadãos. No entanto, os desafios metodológicos na avaliação de situações particulares e os instrumentos de medição actualmente disponíveis tornam necessário aprofundar o conhecimento e melhorar as diferentes técnicas existentes de determinação do custo-utilidade, bem como explorar técnicas alternativas de elicitación de preferências para apoio à tomada de decisão. O EQ-5D é um instrumento genérico de medição em saúde baseado em preferências cujos valores de utilidade são obtidos com recurso à equivalência em tempo. É geralmente reconhecido que a versão tradicional do EQ-5D (EQ-5D-3L) sofre de um efeito-tecto, dificultando a descrição dos estados de saúde situados na parte superior da escala de utilidade. Neste âmbito, o grupo EuroQol expandiu o sistema descritivo deste instrumento de medição, aumentando-o para cinco níveis (EQ-5D-5L). Estudos anteriores demonstram ainda que os valores de utilidade elicitados pela técnica de equivalência em tempo geram pouco consenso e que a utilização desta técnica se revela particularmente complexa na aplicação aos respondentes, para além de excluir grupos particulares como crianças, idosos ou culturas diferentes. A limitação desta técnica tem conduzido a um interesse crescente na experiência de escolha discreta, particularmente em estudos para elicitación de preferências em saúde. Assim, surge este trabalho, integrado numa equipa multidisciplinar internacional liderada pelo grupo EuroQol, que tem por objectivo elicitare as preferências da população portuguesa de forma a estimar o sistema de valores para os 3.125 estados de saúde gerados pelo EQ-5D-5L, com recurso à valoração da equivalência em tempo e da

experiência de escolha discreta. No final deste trabalho espera dispor-se, para a população portuguesa, de um instrumento mais adequado para a medição das preferências em saúde que possa contribuir para o processo de tomada de decisão em políticas de saúde.

Palavras-chave: experiência de escolha discreta, equivalência em tempo, preferências, valoração de estados de saúde, EQ-5D-5L

ABSTRACT

The decision making process within a context of limited resources has important consequences on the health of individuals. The criteria by which health technology assessment and health interventions are decided and the review of current procedures represent a challenge for decision makers in the face of the current economic slowdown. The decision making process requires a transparent and rational approach to ensure the maximization of health and the reduction of inequalities. Economic evaluation techniques may provide a valid solution to achieve these goals. Preference-based measures are part of economic evaluation techniques. These approaches incorporate utilities for health outcomes and can be used in cost-utility analyses to aid resource allocation decisions. They may also meet the needs of citizens. However, the methodological challenges regarding particular situations and the preference based-measures currently in use require further research in order to improve cost-utility analysis and explore alternative preference elicitation methods for decision making. The EQ-5D is a generic preference based measure of health whose value set is obtained using a valuation technique such as time trade-off. It is generally recognized that the traditional version of the EQ-5D (EQ-5D-3L) has a ceiling effect, making difficult to distinguish between the health states situated at the top of the scale. In this context, the EuroQol Group expanded the descriptive system of this generic preference based measure of health, up to five levels (EQ-5D-5L). Previous studies have shown that the utility values elicited by the time trade-off technique generate little consensus and that the use of this technique has proved particularly complex in the application to the respondents and exclude particular groups such as children, the elderly or different cultures. The limitation of this technique has led to a growing interest in the use of the discrete choice experiment, particularly in studies for elicitation of preferences in health. This work, as part of a multidisciplinary international team led by the EuroQol group, aims to elicit the preferences of the Portuguese population for the 3,125 health states generated by the descriptive system of the EQ-5D-5L in order to estimate its value set, using the valuation technique of time trade-off and a discrete choice experiment. In

the end, it expects to provide the Portuguese population with a more appropriate instrument for the measurement of preferences in health that can contribute to the decision-making process in health policies.

Keywords: discrete choice experiment, time trade-off, preferences, health state valuation, EQ-5D-5L

SIGLAS

AIC	<i>Akaike Information Criterion</i>
AMU	Áreas medianamente urbanas
APR	Áreas predominantemente rurais
APU	Áreas predominantemente urbanas
BIC	<i>Bayesian Information Criterion</i>
CAPi	<i>Computer assisted personal interviews</i>
CEISUC	Centro de Estudos e Investigação em Saúde da Universidade de Coimbra
CNPD	Comissão Nacional de Protecção de Dados
cTTO	<i>Composite Time Trade-Off</i>
DALY	<i>Disability Adjusted Life-Year</i>
DCE	<i>Discrete Choice Experiment</i>
EQ-VT	<i>EuroQol Group Valuation Technology</i>
EQ-5D-5L	<i>Five level EuroQol five dimensions questionnaire</i>
GLS	<i>Generalized Least Squares</i>
HUI	<i>Health Utilities Index</i>
HYE	<i>Healthy Years Equivalent</i>
INFARMED	Autoridade Nacional do Medicamento e Produtos de Saúde
MVH	<i>Measurement and Valuation of Health</i>
NICE	<i>National Institute of Health and Clinical Excellence</i>
OLS	<i>Ordinary Least Squares</i>
PIB	Produto Interno Bruto
QALY	<i>Quality-Adjusted Life-Years</i>
QC	<i>Quality Control</i>
QWB	<i>Quality of Well-Being</i>
SAVE	<i>Saved-Young-Life Equivalent</i>
SG	<i>Standard Gamble</i>
TIPAU	Tipologia de Áreas Urbanas

TTO *Time Trade-Off*

VAS *Visual Analogue Scale*

TABELAS

Tabela 1 - Projecção da amostra representativa da população portuguesa, estratificada por género e grupo etário	75
Tabela 2 - Exemplo de quantificação dos padrões de resposta da DCE	104
Tabela 3 - Controlo de qualidade das entrevistas piloto por entrevistador e total de entrevistas	120
Tabela 4 - Distribuição dos entrevistadores por centros de recrutamento.....	123
Tabela 5 - Distribuição da amostra projectada pelo Censos de 2011 e da amostra real do estudo, por género e grupo etário.....	126
Tabela 6 - Características sociodemográficas da amostra	128
Tabela 7 - Estado de saúde/Experiência com doença do entrevistado	131
Tabela 8 - Frequência relativa dos problemas por dimensão e grupo etário	132
Tabela 9 - Frequência absoluta e relativa de indivíduos com e sem problemas, por dimensão e grupo etário	133
Tabela 10 - Estado de saúde global auto-reportado por grupo etário	134
Tabela 11 – Estatísticas descritivas da VAS por grupo etário	135
Tabela 12 - Estatísticas da VAS por pontuação obtida nos níveis de problemas do estado de saúde auto-reportado.....	135
Tabela 13 - Número e percentagem de indivíduos que valorou todos os estados de saúde de forma igual e de funnies	137
Tabela 14 - Controlo de qualidade das entrevistas válidas (n=1.451)	138
Tabela 15 - Duração média do exemplo da cadeira de rodas (n=1.451)	139
Tabela 16 - Número médio de movimentos no exemplo da cadeira de rodas (n=1.451)	142
Tabela 17 - Estados de saúde assinalados no módulo de feedback (n=1.451).....	144
Tabela 18 - Estatísticas descritivas dos valores cTTO.....	146
Tabela 19 - Valores médios e desvio-padrão por índice de intensidade de problemas dos estados de saúde valorados pelo cTTO	147
Tabela 20 - Resultados e avaliação dos modelos OLS, GLS efeitos variáveis e GLS censurado (n=13.391).....	149
Tabela 21 - Resultados e avaliação dos modelos Tobit censurado, Intervalo GLS censurado em 1 (GLS_1) e Intervalo GLS censurado em 1 e 0 (GLS_1_0) (n=13.391)	150
Tabela 22 - Resultados da análise de sensibilidade (modelos para dados cTTO).....	152
Tabela 23 - Quantificação dos padrões de resposta pouco usuais da DCE (n=1.451)	153
Tabela 24 - Resultados e avaliação dos modelos Logit condicional, Logit e Probit (n=10.157).....	155
Tabela 25 - Resultados do modelo híbrido heteroscedástico (Modelo 10).....	158

QUADROS

Quadro 1 - Axiomas da teoria da utilidade de von Neumann-Morgenstern	27
Quadro 2 - Técnicas de medição de preferências.....	28
Quadro 3 - Principais características dos protocolos MVH, Paris e EQ-VT	42
Quadro 4 - Tipo de enviesamento e efeito por técnica de elicitação de preferências.....	45
Quadro 5 - Diferentes abordagens para a obtenção de preferências	65

FIGURAS

Figura 1 - Exemplo do conceito de QALY.....	15
Figura 2 - Aferição de Valor	26
Figura 3 - Jogo-padrão para um estado de saúde crónico preferível à morte	31
Figura 4 - Quadro com roda de probabilidade para aplicação do SG.....	32
Figura 5 – TTO: estado de saúde melhor que morte.....	34
Figura 6 – TTO: estado de saúde pior que morte.....	35
Figura 7 - Lead-time TTO: estado de saúde melhor que morte	36
Figura 8 - Lead-time TTO: estado de saúde pior que morte	37
Figura 9 - Lead-time e Lag-time TTO	38
Figura 10 - cTTO. a) TTO convencional para valoração de estados de saúde melhores que morte	39
Figura 11 - cTTO. b) Lead-time TTO para valoração de estados de saúde piores que morte	40
Figura 12 - Esquemas de iteração dos protocolos de valoração do EuroQol.....	44
Figura 13 - Exemplo de aplicação de um par de cenários para escolha discreta	48
Figura 14 - Versão portuguesa do questionário EQ-5D-5L.....	62
Figura 15 - Países com estudos de valoração do EQ-5D-5L concluídos ou em curso (Fevereiro de 2017).....	72
Figura 16 - EQ-VT: hierarquia de perfis e respectivas funções do utilizador	73
Figura 17 - Acesso ao EQ-VT	80
Figura 18 - Carregamento de questionários após registo do participante.....	81
Figura 19 - Boas-vindas e apresentação do estudo.....	82
Figura 20 - EQ-5D-5L: avaliação actual da mobilidade	83
Figura 21 - EQ-5D-5L: avaliação actual dos cuidados pessoais	83
Figura 22 - EQ-5D-5L: avaliação actual das actividades habituais.....	84
Figura 23 - EQ-5D-5L: avaliação actual da dor/mal-estar.....	84
Figura 24 - EQ-5D-5L: avaliação actual da ansiedade/depressão	85
Figura 25 - VAS.....	85
Figura 26 - Perguntas de contexto pessoal	86
Figura 27 - Introdução ao exemplo da cadeira de rodas.....	87
Figura 28 - Exemplo da cadeira de rodas (início da iteração): estado de saúde melhor que morte (TTO convencional).....	88
Figura 29 - Exemplo da cadeira de rodas (final da iteração): estado de saúde melhor que morte (TTO convencional).....	88
Figura 30 - Exemplo da cadeira de rodas: estado de saúde considerado pior que morte.....	89
Figura 31 - Valoração cTTO: estados de saúde de treino	91
Figura 32 - Valoração de estado de saúde com recurso ao cTTO	93
Figura 33 - Avaliação da dificuldade da tarefa cTTO	94
Figura 34 - Módulo de feedback.....	95
Figura 35 - Caixa de comentários associada à tarefa cTTO	96
Figura 36 - Apresentação da tarefa de valoração DCE	96
Figura 37 - Valoração de um par de estados de saúde com recurso à DCE	97
Figura 38 - Avaliação da dificuldade da tarefa DCE	98
Figura 39 - Caixa de comentários associada à tarefa DCE	99
Figura 40 - Finalização do questionário principal e agradecimentos	99

Figura 41 - QC Tool.....	101
Figura 42 - Distribuição da amostra por concelho de residência do entrevistado	130
Figura 43 - Exemplo de investigação futura no desenho da DCE.....	174

GRÁFICOS

Gráfico 1 - Exemplo da distribuição dos valores cTTO.....	103
Gráfico 2 - Período de recolha de dados associado a cada entrevistador activo	123
Gráfico 3 - Comparação dos grupos por grupo etário e género: amostra projectada pelo Censos de 2011 versus estudo	127
Gráfico 4 - Percentagem de problemas reportados por grupo etário	134
Gráfico 5 - Valores médios VAS por género e grupo etário	136
Gráfico 6 - Relevância das dimensões do sistema descritivo do EQ-5D-5L	137
Gráfico 7 - Duração média do exemplo da cadeira de rodas, por entrevistador	140
Gráfico 8 - Duração média da explicação do estado de saúde melhor que morte no exemplo da cadeira de rodas, por entrevistador	141
Gráfico 9 - Duração média da explicação do estado de saúde pior que morte no exemplo da cadeira de rodas, por entrevistador	141
Gráfico 10 - Número médio de movimentos no exemplo da cadeira de rodas, por entrevistador	142
Gráfico 11 – Número médio de movimentos no estado de saúde melhor que morte no exemplo da cadeira de rodas, por entrevistador	143
Gráfico 12 - Número médio de movimentos no estado de saúde pior que morte no exemplo da cadeira de rodas, por entrevistador	143
Gráfico 13 - Número de participantes com inconsistências relativamente ao número total de entrevistas.....	144
Gráfico 14 - Percentagem de participantes com inconsistências antes e depois da utilização do módulo de feedback	145
Gráfico 15 - Distribuição de valores cTTO por entrevistador.....	146
Gráfico 16 - Percentagem de participantes que escolheu a Vida A ou a Vida B na tarefa DCE comparada com a intensidade relativa de problemas de A e B	154
Gráfico 17 - Valores médios e desvio-padrão por pontuação obtida nos níveis de problemas dos estados de saúde valorados pelo cTTO.....	157
Gráfico 18 - Valor dos estados de saúde por nível e dimensão do EQ-5D-5L.....	159
Gráfico 19 - Comparação dos valores estimados para os 3.125 estados de saúde gerados pelo EQ-5D-5L com recurso ao modelo de regressão por intervalo GLS censurado em 1 e ao modelo híbrido heteroscedástico	160
Gráfico 20 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Holanda (modelo cTTO).....	162
Gráfico 21 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Uruguai (modelo cTTO).....	162
Gráfico 22 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Inglaterra (modelo híbrido)	163
Gráfico 23 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Japão (modelo híbrido)	164

SUMÁRIO

INTRODUÇÃO.....	1
1. Avaliação Económica em Saúde.....	5
1.1. A avaliação económica no processo de tomada de decisão em saúde.....	5
1.2. Técnicas de avaliação económica.....	7
1.2.1. Análise Custo-Benefício.....	8
1.2.2. Análise Custo-Efectividade.....	9
1.2.3. Análise Custo-Utilidade.....	10
2. MEDIDAS DE AVALIAÇÃO ECONÓMICA.....	13
2.1. O QALY como medida de avaliação económica.....	13
2.2. Unidades de medida alternativas ao QALY.....	21
2.2.1. Anos saudáveis equivalentes.....	21
2.2.2. Anos salvos equivalentes de vida jovem.....	23
2.2.3. Anos de vida ajustados por incapacidade.....	23
3. MEDIÇÃO DE PREFERÊNCIAS EM SAÚDE.....	25
3.1. Visual Analogue Scale.....	29
3.2. Standard Gamble.....	30
3.3. Time Trade-Off.....	32
3.4. Limitações das técnicas de medição de preferências mais comuns.....	45
3.5. Discrete Choice Experiment.....	47
4. MEDIDAS DESCRITIVAS DE RESULTADOS EM SAÚDE.....	57
4.1. Quality of Well-Being.....	58
4.2. Health Utilities Index.....	58
4.3. 15-D.....	59
4.4. EQ-5D.....	59
4.5. SF-6D.....	62
4.6. Fontes de Informação.....	64
5. MATERIAIS E MÉTODOS.....	67
5.1. Objectivo.....	67
5.2. Questão de Investigação.....	68
5.3. Procedimentos.....	68
5.3.1. Apresentação da candidatura.....	68
5.3.2. Código de conduta.....	69

5.3.3. Planeamento.....	71
5.3.4. Amostra e recolha de dados.....	73
5.3.5. Preceitos ético-legais.....	79
5.4. Instrumento	79
5.4.1. Questionário local.....	100
5.5. Controlo de qualidade.....	100
5.5.1. Ferramenta de controlo de qualidade: QC Tool	100
5.5.2. Reinquirição	104
5.6. Análise estatística	105
5.6.1. Determinação dos blocos de estados de saúde.....	105
5.6.2. Critérios de exclusão de dados.....	107
5.6.3. Modelação cTTO	108
5.6.4. Modelação DCE.....	113
5.6.5. Modelação híbrida.....	116
5.6.6. Análise de sensibilidade	117
6. RESULTADOS	119
6.1. Supervisão da recolha de dados.....	119
6.2. Caracterização da amostra	126
6.2.1. Representatividade.....	126
6.2.2. Análise descritiva da amostra	127
6.3. Modelo cTTO.....	138
6.3.1. Controlo de qualidade	138
6.3.2 Análise dos dados	146
6.4. Modelo DCE.....	153
6.4.1. Controlo de qualidade	153
6.4.2. Análise dos dados	153
6.5. Modelo híbrido	156
6.6. Comparação dos sistemas de valores, cTTO e híbrido, obtidos para Portugal com outros sistemas de valores internacionais	161
7. DISCUSSÃO	165
8. CONCLUSÃO	181
REFERÊNCIAS BIBLIOGRÁFICAS	183
APÊNDICE I	203

INTRODUÇÃO

Nas políticas de saúde, tal como em outras áreas, são tomadas decisões tendo por base a existência de recursos necessariamente sempre limitados. Em quaisquer dos sistemas de saúde, estas decisões têm consequências claras na saúde das populações.

Tal como defendido por Baltussen e Niessen (2006), para garantir um nível óptimo na utilização dos recursos, o estabelecimento de prioridades na tomada de decisão não pode ser um processo *ad hoc*.

Neste âmbito, surge a avaliação económica como uma forma de decidir sobre a correcta alocação de recursos (financeiros, humanos, tecnológicos, formativos,...) e que tem como principal objectivo o apoio aos agentes de decisão na escolha de opções eficientes e equitativas, quando comparando custos e benefícios de programas e intervenções alternativas em saúde (Brazier et al. 2007; Drummond et al. 2005).

O processo não se resume apenas a decidir sobre diferentes alternativas de acordo com os ganhos e custos em saúde que estas originam. A tomada de decisão é bem mais complexa e inclui múltiplos critérios como a equidade social, a qualidade da experiência do utilizador do serviço, o impacto na economia em geral e a evidência científica (Devlin & Sussex 2011). O processo de escolha de tecnologias e intervenções em saúde, que conquista cada vez mais terreno nos países desenvolvidos e em desenvolvimento, tem contribuído para uma maior eficiência na alocação de recursos. Os múltiplos desafios a que os sistemas públicos de saúde estão sujeitos quando procuram a maximização do estado de saúde da população e a redução das desigualdades em saúde, no actual contexto económico e de pressão substancial nos orçamentos do Estado, revestem a tomada de decisão de particular complexidade face à escolha de novas tecnologias ou programas de saúde e na reavaliação de procedimentos já existentes. Baltussen e Niessen (2006) referem que os decisores gerem com alguma dificuldade problemas desta natureza, recorrendo a abordagens heurísticas ou intuitivas para diminuir a sua complexidade. Os decisores também podem estar sujeitos a motivações individuais e a pressões políticas que vão constituir um factor perturbador. O risco de arbitrariedade na tomada de decisão é, por

isso, real. A variabilidade da decisão em situações aparentemente semelhantes pode, por vezes, ser explicada por questões contextuais, tais como: valores pessoais dos decisores, distintas percepções de risco, diferentes prioridades e disponibilidade orçamental variável. Por esse motivo, a escolha do método mais adequado para decidir sobre a tecnologia, programa ou intervenção em saúde reveste-se de extrema importância, nomeadamente no que concerne à independência, transparência e isenção da definição e medição de valor. A preocupação pela utilização de abordagens racionais e transparentes é também partilhada por autores como Baltussen e Niessen (2006) e Devlin e Sussex (2011).

A avaliação económica tem por objectivo identificar, medir, valorar e comparar custos e consequências quando a escolha de diversas alternativas está em causa, assumindo que a escolha de uma opção em detrimento de outra(s) conduz a uma diferente forma de utilização dos recursos (Brazier et al. 2007). Para uma decisão informada, Drummond et al. (2005) enfatizam a necessidade de realizar uma análise sistemática para a identificação clara de todas as alternativas relevantes, a importância de conhecer as diversas perspectivas dos intervenientes envolvidos e da determinação dos custos de oportunidade associados, comparando-os com os benefícios obtidos em cada alternativa. Em teoria, Devlin e Sussex (2011) defendem que, se fosse possível determinar o custo-efectividade de todos os serviços e tecnologias de saúde, seria possível maximizar o benefício da população sujeito a um orçamento específico. Para tal, não só se torna necessário, independentemente do cenário envolvido, determinar o custo-efectividade com a maior exactidão possível mas também conhecer as perspectivas dos intervenientes e as suas preferências. Dados os desafios metodológicos na avaliação de situações particulares (crianças, idosos, patologias específicas) é necessário aprofundar o conhecimento e melhorar as diferentes técnicas existentes de determinação do custo-efectividade bem como explorar técnicas alternativas de elicitación de preferências para apoio à tomada de decisão. Neste sentido, a principal motivação para a realização deste trabalho é contribuir para a obtenção de instrumentos de medição em qualidade em saúde que melhor representem as preferências dos cidadãos.

Esta tese de doutoramento encontra-se inserida num projecto de investigação internacional, liderado pelo grupo EuroQol, que integra uma rede multilingue e

multidisciplinar de investigadores. O projecto internacional assume como principal objectivo obter os sistemas de valores para os estados de saúde gerados pelo instrumento genérico de medição em saúde EQ-5D-5L nos diversos países.¹ A necessidade de obter os diferentes sistemas de valores decorre de ser comumente aceite que as preferências das populações diferem de país para país (Huang et al. 2007; Johnson et al. 1998; Hawthorne et al. 2007; Ferreira & Ferreira 2014). Neste sentido e sob a coordenação do grupo EuroQol, foram já finalizados seis estudos de valoração a nível internacional com o intuito de obter o sistema de valores para os estados de saúde definidos pelo EQ-5D-5L, estando em curso, embora em estádios diferentes, vários outros estudos em diversos países. A nível nacional este projecto é coordenado pelo Centro de Estudos e Investigação em Saúde da Universidade de Coimbra (CEISUC). Neste contexto surge esta tese de doutoramento que tem por objectivo a estimação do sistema de valores da população portuguesa para os diferentes estados de saúde descritos pelo instrumento de medição em saúde EQ-5D na sua versão mais recente (EQ-5D-5L), já utilizada previamente na população portuguesa no estudo efectuado por Ferreira e colaboradores (2016), com recurso a duas técnicas distintas de elicitación de preferências: a equivalência em tempo e a experiência de escolha discreta. O grupo coordenador nacional deste projecto é constituído por um investigador principal e três co-investigadores: Pedro Lopes Ferreira, que integra o CEISUC e a Faculdade de Economia da Universidade de Coimbra (investigador principal), Lara Noronha Ferreira, pertencente ao CEISUC e à Escola Superior de Gestão, Hotelaria e Turismo da Universidade do Algarve, Luís Pereira, do Centro de Investigação sobre o Espaço e as Organizações da Universidade do Algarve e Escola Superior de Gestão, Hotelaria e Turismo da Universidade do Algarve, e Patrícia Antunes, aluna de doutoramento que integra o CEISUC.

Este trabalho encontra-se estruturado em três partes principais. Na primeira parte (capítulos 1 a 4) efectua-se uma revisão do estado da arte relativamente ao tema abordado. A segunda parte (capítulo 5) incide sobre os materiais e métodos utilizados na condução deste estudo. Por fim, a última parte consubstanciada nos capítulos 6, 7 e 8 apresenta os resultados, a discussão, as conclusões e ainda as referências bibliográficas.

¹ EQ-5D-5L – *Five level EuroQol five dimensions questionnaire*.

1. Avaliação Económica em Saúde

A afectação óptima dos recursos escassos na saúde constitui uma parte importante no processo de tomada de decisão. A avaliação económica, com foco primário na obtenção de eficiência ou do máximo benefício dados os recursos disponíveis, tem sido desenvolvida para apoio à decisão.² Se os recursos fossem ilimitados e os sistemas de saúde tivessem a possibilidade de proporcionar qualquer tecnologia de saúde aos cidadãos que dela necessitassem, nesse caso, não haveria necessidade de se recorrer à avaliação económica. No entanto, dado os recursos serem escassos, esta situação torna-se irreal, conduzindo inevitavelmente à escolha de apenas algumas opções em detrimento de outras.

Neste âmbito, a avaliação económica em saúde torna-se uma ferramenta relevante que permite a definição do valor obtido em saúde através do custo que pressupõe a opção por uma tecnologia ou intervenção em saúde. A sua utilização crescente para apoio à decisão tem promovido o estabelecimento de requisitos rigorosos no que concerne aos métodos de análise, nomeadamente a definição do problema de decisão, a adopção de um horizonte temporal adequado para a investigação em curso, a especificação de uma função objectivo sob a forma de uma medida de ganhos em saúde, a inclusão de uma perspectiva de custos, a necessidade de incorporar na análise toda a evidência relevante e a inclusão de incerteza (Briggs et al. 2006; Sculpher et al. 2005).

1.1. A avaliação económica no processo de tomada de decisão em saúde

Tal como sucede noutros países desenvolvidos, também em Portugal uma parte significativa do Produto Interno Bruto (PIB) é alocada aos serviços de saúde. Em 2015, o valor previsional alocado, em relação ao PIB, foi de 8,9% (EU & OECD 2016). O valor

² Em inglês, o conceito refere-se a “*value for money*”.

destinado à saúde, quando comparado com outros bens e serviços, depende do crescimento dos gastos em saúde bem como do crescimento da economia. Não obstante o crescimento verificado em Portugal, desde 2005 até 2009, nos gastos em saúde em relação ao PIB, a partir desse ano verificou-se uma variação negativa, possivelmente associada à crise económica que o país atravessou (EU & OECD 2016). Noutro contexto económico seria expectável que factores como um maior envelhecimento populacional, o aumento da prevalência de doenças crónicas, a ausência de medidas de contenção de custos e a inovação tecnológica induzissem o crescimento dos gastos em saúde relativamente ao PIB. Destes, a inovação tecnológica é o principal factor para o crescimento dos gastos a longo prazo em cuidados de saúde (Roseboom et al. 2017). Considerando o previsível aumento nos gastos em saúde a um ritmo superior ao do crescimento económico e a consequente pressão nos orçamentos públicos, infere-se que o processo de tomada de decisão reveste-se de cada vez maior importância.

A avaliação económica é utilizada no processo de tomada de decisão em saúde por diversas razões que, de acordo com Morris, Devlin & Parkin (2007) poderão ser elencadas como i) maximizar os benefícios decorrentes dos gastos em saúde; ii) ultrapassar as desigualdades regionais no acesso à saúde; iii) limitar os custos e gerir a procura; e iv) providenciar poder negocial junto dos fornecedores de produtos e tecnologias de saúde. O motivo mais comum para a sua utilização é o de conter custos e gerir a procura de cuidados, materializado pela avaliação de tecnologias em saúde como forma de controlo da introdução de novas tecnologias habitualmente com custo muito elevado. Não obstante a avaliação de tecnologias ser dominada pela indústria farmacêutica, os governos de diversos países têm também adoptado técnicas de avaliação económica para apoio à decisão no reembolso de custos aos doentes ou prestadores de serviços (por exemplo, no que concerne à utilização de medicamentos). São várias as organizações, em distintos países, que utilizam técnicas de avaliação económica para informar os decisores. Destas, destaca-se a Autoridade Nacional do Medicamento e Produtos de Saúde (INFARMED) em Portugal, o *National Institute of Health and Clinical Excellence* (NICE), o *Scottish Medicines Consortium* (Escócia) e o *All Wales Medicines Strategy Group* (Gales) no Reino Unido, o *National Centre for Pharmacoeconomics* e a *Health Information and Quality Authority* na

Irlanda, em França a *Haute Autorité de Santé*, em Espanha a agência nacional *Instituto de Salud Carlos III* em colaboração com mais sete agências regionais, a *Food and Drug Administration* nos Estados Unidos da América e a *Canadian Agency for Drugs and Technologies in Health* no Canadá, entre outros.

1.2. Técnicas de avaliação económica

Drummond e colaboradores (2005) identificam várias técnicas de avaliação económica que permitem uma avaliação comparativa de custos e consequências de diferentes alternativas. Essas técnicas são: i) análise custo-benefício; ii) análise custo-efectividade; e iii) análise custo-utilidade.³

Perante todas as opções existentes de técnicas de avaliação económica, a escolha entre as diversas formas depende não só da perspectiva do analista em termos de avaliação económica mas também de vários factores tais como o problema em apreço, os dados disponíveis para análise, os desafios associados às unidades de medida e as limitações que cada técnica incorpora.

No âmbito do apoio à decisão, o analista pode assumir uma perspectiva em que considera que a melhor forma de proceder à avaliação económica é determinar o valor que o indivíduo está disposto a pagar pelas alternativas em análise, de forma a comparar com os custos inerentes e determinar qual a opção mais vantajosa. Esta é uma perspectiva no âmbito da economia do bem-estar (em inglês, abordagem *welfarist*) que coloca um enfoque significativo no valor que os indivíduos associam aos resultados, considerando os indivíduos como os melhores avaliadores do seu próprio bem-estar. Neste contexto, insere-se a análise custo-benefício. No entanto, esta técnica não é isenta de críticas na medida em que o valor que o indivíduo expressa estar disposto a pagar pode ser condicionado pela sua disponibilidade financeira e pela noção individual que este possui do valor das alternativas em análise (Drummond et al. 2005). Por outro lado, o analista

³ Os conceitos subjacentes a estas técnicas, definidos inicialmente por Drummond et al. (2005), variam de autor para autor, tendo sofrido alguma evolução ao longo do tempo. No que concerne a esta tese, adoptam-se as definições apresentadas por Drummond et al. (2005).

pode entender a avaliação económica como uma forma de melhor alocar os recursos financeiros em saúde. Neste caso, a análise efectuada incide apenas sobre atributos de saúde (isto é, exclui atributos não relacionados com a saúde como a disponibilidade para pagar), comparando os recursos de saúde consumidos com os ganhos em saúde obtidos, medidos, por exemplo, em unidades de medida de saúde ou efeitos de saúde, embora seja preferível a utilização de índices de forma a efectuar comparações transversais entre programas. Esta é a perspectiva mais comum entre os decisores na área da saúde e enquadra-se no âmbito de uma perspectiva *extrawelfarist*, inserindo-se, neste contexto, as análises custo-efectividade e custo-utilidade (Drummond et al. 2005).

De uma forma sucinta, passa-se a apresentar de seguida as diferentes técnicas de avaliação económica elencadas.

1.2.1. Análise Custo-Benefício

A análise custo-benefício avalia custos e consequências de diferentes alternativas e, para facilitar a comparação entre elas, utiliza uma unidade de medida monetária. Para além de medir, este tipo de análise atribui um valor monetário a todos os resultados de uma intervenção, isto é, valora os resultados em saúde em termos de disponibilidade para pagar dos indivíduos. A avaliação custo-benefício determina o valor monetário resultante da diferença entre os custos envolvidos na utilização de uma alternativa com os ganhos/poupança/benefícios obtidos com a opção da sua utilização. Na prática, pressupõe-se que, na análise custo-benefício, uma alternativa não poderá ser escolhida se o seu benefício líquido for negativo. Tal não sucede nas análises custo-efectividade e custo-utilidade, particularmente na área da saúde, em que se assume que se um determinado resultado proporciona ganhos em saúde, importa apenas escolher a alternativa que o oferece com maior custo-efectividade (i.e., que possibilita o mesmo resultado com menor custo). Isso significa que, neste tipo de análises, poderá estar em causa a escolha de alternativas que não são auto-sustentáveis (Drummond et al. 2005).

Para decidir sobre a alocação de recursos recorrendo às análises custo-efectividade e custo-utilidade, como estas técnicas não atribuem um valor monetário aos

resultados em saúde, torna-se necessário utilizar critérios externos associados a um valor frequentemente considerado arbitrário (e.g. limites oficiais publicados, tabelas de classificação de custo). Drummond e colaboradores (2005) consideram que a análise custo-benefício, ao incluir como parte da análise o valor monetário dos resultados em saúde, pode constituir-se como mais interessante relativamente a outras técnicas, ao possibilitar a sua utilização mais abrangente, i.e., alargando o seu uso a outros sectores da economia para além do da saúde (Drummond et al. 2005). Não deixa de ser, no entanto, relevante que o conceito de disponibilidade para pagar possa ser limitado pela capacidade de pagar, a qual depende directamente do rendimento disponível e do valor que cada indivíduo atribui ao dinheiro e às alternativas em análise.

1.2.2. Análise Custo-Efectividade

A análise custo-efectividade compara os custos e os resultados de efectividade de várias alternativas que concorrem para o mesmo objectivo. Quando duas ou mais alternativas atingem o mesmo nível do objectivo determinado, a alternativa com menor custo é considerada como a mais custo-efectiva. Neste tipo de avaliação económica, a unidade de medida utilizada é o custo por unidade de efeito pretendido (e.g. número de anos de vida ganhos, número de casos de cancro detectados).

Quando os resultados de diferentes alternativas são praticamente equivalentes, a análise custo-efectividade poderá ser considerada como uma análise de minimização de custo, embora esta seja uma perspectiva pouco habitual devido à incerteza e variabilidade associadas à estimação dos custos e dos resultados.

A análise custo-efectividade é particularmente utilizada nos casos em que o decisor, limitado a um orçamento específico, considera um conjunto limitado de opções, numa determinada área de actuação. Tome-se, por exemplo, o caso em que um decisor é responsável por organizar um programa de rastreio oncológico e onde, como resultado, poderá estar interessado no número de casos de cancro detectados. No entanto, é facilmente perceptível que o benefício obtido por se detectar casos de cancro é variável em função do tipo de cancro, da idade do doente e da fase de desenvolvimento da doença. Se se considerar que o objectivo principal dos programas de rastreio oncológico é prevenir

a morbilidade e mortalidade prematuras, o número de casos detectados poderá não ser uma unidade de medida ou efeito desejável (Drummond et al. 2005).

Na verdade, a análise custo-efectividade apresenta várias limitações. Pelo facto das unidades de efeito diferirem de intervenção para intervenção, esta técnica não pode ser utilizada para efectuar comparações entre diferentes programas ou intervenções. Por outro lado, este tipo de avaliação económica não permite obter a informação relativa ao custo de oportunidade de financiar uma nova intervenção que se revele custo-efectiva. Ao utilizar apenas uma unidade de efeito ou resultado único, limita a avaliação de intervenções com mais do que um tipo de resultado de interesse. Por fim, não questiona nem valoriza os objectivos da medida a avaliar.

Neste sentido, a análise custo-efectividade é considerada uma técnica restritiva, impossibilitando, assim, uma decisão informada sobre a eficiente alocação de recursos (Brazier et al. 2007; Drummond et al. 2005).

1.2.3. Análise Custo-Utilidade

A análise custo-utilidade foi desenvolvida para responder às limitações da análise custo-efectividade.

Na análise custo-utilidade, o conceito de utilidade é tido como a preferência que um indivíduo ou uma sociedade possui por um determinado conjunto de resultados em saúde.

Drummond et al. (2005) consideram a análise custo-utilidade similar à análise custo-efectividade no que concerne aos custos, mas diferente no que respeita aos tipos de resultados ou efeitos. Enquanto que na análise custo-efectividade os resultados são únicos, específicos de um programa e não valorados, na análise custo-utilidade os resultados podem ser únicos ou múltiplos, genéricos (i.e. não relativos a um programa específico) e incorporar a noção de valor associado aos resultados. Por esse motivo, a análise custo-utilidade possui uma utilização mais abrangente, tornando-se mais útil para os decisores. Neste tipo de avaliação económica, o efeito respeita ao número de anos com saúde perfeita e, por isso, a unidade de medida mais comumente utilizada é o ano de vida ajustado à qualidade (*Quality-Adjusted Life-Years* ou QALY) que incorpora, num só

indicador, a quantidade e a qualidade dos anos de vida de um indivíduo. Os mesmos autores referem outras alternativas genéricas para medidas de resultados como os anos saudáveis equivalentes (*Healthy Years Equivalent* ou HYE), os anos de vida ajustados por incapacidade (*Disability Adjusted Life-Year* ou DALY) e os anos salvos equivalentes de vida jovem (*Saved-Young-Life Equivalent* ou SAVE).⁴

Uma característica distintiva da análise custo-utilidade é medir a preferência relativa pelos resultados de acordo com a teoria de utilidade de von Neumann-Morgenstern (Drummond et al. 2005).

Na análise custo-utilidade a unidade de efeito é o custo por cada ano saudável ou o custo por cada QALY ganho por se escolher um determinado tratamento ou programa em detrimento de outro. Neste tipo de avaliação económica o recurso aos QALY possibilita diversas vantagens porque condensa num só valor toda a análise multidimensional e, por isso, permite comparar alternativas com vários tipos de resultados em saúde. No entanto, essa agregação simples faz com que um QALY tenha sempre o mesmo valor, independentemente de quem o experiencia.

⁴ A maior parte da literatura refere-se à terminologia destas unidades de medida em inglês. Por esse motivo, ao longo do texto, serão utilizadas as respectivas siglas em inglês.

2. MEDIDAS DE AVALIAÇÃO ECONÓMICA

Num mercado perfeito, a oferta e a procura de bens ou serviços atingem o equilíbrio, sem regulações externas ou interferências do Estado. No entanto, este resultado é condicionado quando existem falhas de mercado. É o que sucede no mercado da saúde nomeadamente pelo facto de existir incerteza (e.g., no que respeita à alteração do estado de saúde e do tipo de tratamento a escolher), assimetria de informação (e.g., entre prestadores de serviços de saúde e consumidores) e externalidades (e.g., caso da imunidade de grupo que se atinge pelo facto de cada um dos indivíduos se vacinar). A existência de falhas de mercado na saúde, conducentes a ineficiências e desigualdades, justificam a regulação deste mercado, para alcançar a optimização na alocação e na distribuição dos recursos escassos (Brazier et al. 2007). Isto implica a definição de critérios de decisão tendo por base técnicas de avaliação económica que traduzem benefícios para os indivíduos. Tal como referido anteriormente, as diferentes técnicas de avaliação económica diferem pelas unidades que utilizam para valorar os benefícios dos cuidados de saúde. Os subcapítulos seguintes descrevem a medida de avaliação económica de uso mais generalizado na decisão sobre alocação de recursos em saúde – o QALY – e outras unidades alternativas de medida.

2.1. O QALY como medida de avaliação económica

O QALY convencional foi originalmente desenvolvido em avaliação económica como uma medida de efectividade em saúde, com o intuito de apoiar a decisão sobre a alocação de recursos escassos em programas de saúde concorrentes (Weinstein et al. 2009). Trata-se da medida de referência na qual as agências reguladoras e de contratação de vários países tais como INFARMED (Portugal), o *US Panel on Cost-Effectiveness in Health and Medicine* (Estados Unidos da América), o NICE (Reino Unido), a *Haute Autorité de Santé* (França), o *Institut für Pharmakoökonomische Forschung* (Áustria), o *Health Information and*

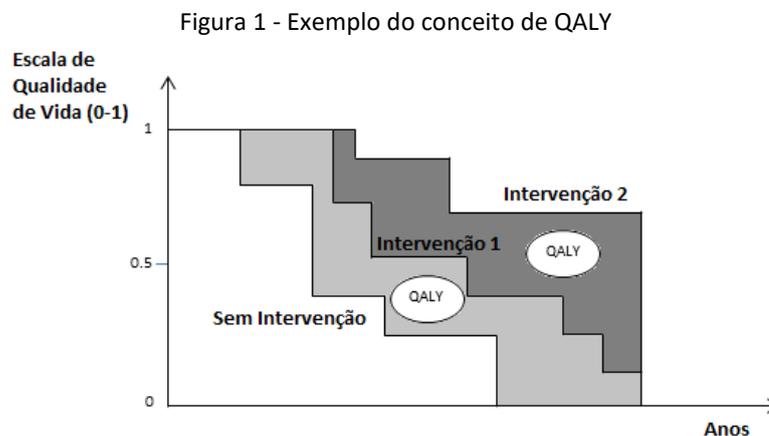
Quality Authority (Irlanda), a *Canadian Agency for Drugs and Technologies in Health* (Canadá), o *College voor zorgverzekeringen* (Holanda), o *Ministry of Social Affairs and Health* (Finlândia) ou o *Statens legemiddelverk* (Noruega) se baseiam para efectuar técnicas de avaliação económica em saúde, reconhecendo o QALY como uma medida válida para avaliações de produtos médicos e de intervenções de saúde (Ferreira & Ferreira 2014; Lipscomb et al. 2009). Esta medida é também utilizada em tabelas de classificação de custo por QALY (em inglês, *QALY league tables*) com o intuito de ordenar programas ou tecnologias de saúde em função da sua eficiência ou do menor rácio custo-efectividade incremental (em inglês, *Incremental Cost-Effectiveness Ratio*). No entanto, para uma utilização adequada destas tabelas, é necessário dispor de dados locais específicos, devidamente contextualizados, o que nem sempre sucede por falta de recursos, tempo ou de conhecimentos técnicos para a condução deste tipo de análise, obrigando a recorrer a tabelas já publicadas em outros estudos que poderão não ser tão relevantes para a análise local (Mason 1994). Por este motivo, há poucos exemplos de tabelas de classificação de custo por QALY utilizadas na priorização de recursos (Williams et al. 2012).

O QALY é uma unidade de medida que constituiu uma verdadeira inovação em economia da saúde, ao combinar numa única medida o benefício dos cuidados de saúde, em termos de longevidade e de qualidade dos anos de vida (Brazier et al. 2007). A missão de otimizar os recursos que a sociedade utiliza, a possibilidade de utilização em situações distintas (e.g., em diferentes patologias) e nos âmbitos individual e populacional torna o QALY uma medida preferencial para apoio à decisão, sobretudo por parte de organismos públicos que recorrem a técnicas de avaliação económica (Kobelt 2013).

Para concretizar o conceito de QALY é necessária a identificação de pesos que representam a qualidade de vida relacionada com a saúde dos estados de saúde sob avaliação. Estes pesos devem ser baseados em preferências, ancorados nos pontos de referência “saúde perfeita” e “morte” e medidos numa escala intervalar (Drummond et al. 2005).

O modelo QALY assume que os indivíduos se movem através de estados de saúde ao longo do tempo e que cada estado de saúde tem um valor associado. De acordo com Brazier et al. (2007), o QALY é um indicador que mede a qualidade de vida relacionada com

a saúde, atribuindo para cada par (período de tempo/estado de saúde) um peso específico, que varia entre 0 (estado equivalente a morte) e 1 (estado equivalente a saúde perfeita).⁵ Os estados de saúde considerados piores que morte assumem valores negativos.⁶ Note-se que esta abordagem valora estados de saúde e não alterações entre estados de saúde. Na prática tem-se que, para um determinado indivíduo, a perspectiva de viver Y anos num estado de saúde inferior ao estado de saúde perfeita pode ser equacionada pela perspectiva de viver X anos em perfeita saúde, considerando $X < Y$ (Brazier et al. 2007). Em termos gráficos, este conceito pode ser representado de acordo com a Figura 1:



Fonte: Adaptado de Cornes, P. (2012). The economic pressures for biosimilar drug use in cancer medicine. *Targeted Oncology*. 7 (Supl 1), S57–S67, Figura 3

No exemplo apresentado na Figura 1 é possível observar o efeito dos QALY de acordo com a intervenção efectuada: a intervenção 1 aumenta a qualidade e a quantidade de anos vividos enquanto a intervenção 2 aumenta apenas a qualidade de vida (o número de anos de vida mantém-se igual ao da intervenção 1). O número de QALY associado a cada intervenção é representado pela área entre cada alternativa (sem intervenção, intervenção 1 e intervenção 2). No caso apresentado na Figura 1, o benefício líquido da intervenção 2 corresponde à área da intervenção 2 subtraída da área correspondente à ausência de

⁵ O par de valores (0, 1) respeita aos pontos de referência na escala convencional do QALY. No entanto, outros pares de valores poderão ser utilizados uma vez que se trata de uma escala de razão ou proporcional (Drummond et al. 2005).

⁶ Esta escala permite ainda que estados de saúde melhores que saúde perfeita, se existissem, fossem cotados com valores superiores a 1 (Drummond et al. 2005).

intervenção e da área referente à intervenção 1. A título de exemplo, se se considerar que um indivíduo que padece de uma determinada patologia vê a sua qualidade de vida reduzida, por esse motivo, para 0,01 durante 15 anos e, se o mesmo indivíduo for submetido a um tratamento para essa patologia que lhe permite alcançar uma qualidade de vida de 0,6 durante 20 anos, os QALY obtidos calculam-se da seguinte forma:

$$(20 \times 0,6) - (15 \times 0,01) = 11,85$$

Considerando uma situação com ausência de incerteza, sem desconto temporal e sem alterações no estado de saúde, o número de QALY relativos a um resultado em saúde calcula-se multiplicando o valor atribuído a um estado de saúde específico pelo número de anos passados nesse estado. Se se admitir que pequenas alterações do estado de saúde ocorrem, quando se passa de um período de tempo para outro, o algoritmo do QALY passa a ser uma combinação de valores de estados de saúde, durações e probabilidades. Assim, o QALY convencional pode ser definido de acordo com o seguinte:

$$QALY_{conv} = \sum_{t=1 \dots T} \sum_{s=1 \dots S} p_{st} V(H_{st})(1+r)^{t-1} \quad [1]$$

em que p_{st} corresponde à probabilidade do indivíduo se encontrar no estado de saúde H_s , no período t ; $V(H_{st})$ é o valor atribuído ao estado de saúde H_s , no período t ; $(1+r)^{t-1}$ respeita a um factor de desconto que serve para atribuir valores correntes a $V(H_{st})$ e em que r corresponde à taxa de desconto temporal; S é o número discreto de estados de saúde que podem ocorrer; e T respeita ao horizonte temporal relevante para a tomada de decisão (Lipscomb et al. 2009). Quando os QALY são usados para decisões ao nível da sociedade, V reflecte habitualmente o valor médio das preferências dos indivíduos que constituem o grupo. Note-se que tanto os indivíduos como a sociedade em geral preferem que os ganhos em saúde ocorram num momento mais próximo do que futuro e, por esse motivo, os valores obtidos são multiplicados por um factor de desconto que justifica essa preferência temporal (Drummond et al. 2005). Este factor de desconto não é, no entanto, livre de crítica

quando se utiliza uma técnica de elicitación de preferências que já incorpora uma preferência temporal (e.g. equivalência em tempo). Não obstante os valores obtidos por técnicas como o jogo-padrão não serem afectados por preferências temporais, o mesmo não sucede com a equivalência em tempo. Além disso, o valor habitualmente recomendado para a taxa social de desconto, pela maior parte das linhas orientadoras de avaliação económica (3 a 5%), pode não representar a preferência temporal do indivíduo. Apesar de assumirem ser necessário aprofundar a investigação sobre estes constrangimentos, Drummond et al. (2005) continuam a recomendar a prática de desconto temporal, no valor recomendado da taxa social, independentemente da técnica de elicitación de preferências utilizada.

Brazier et al. (2007) explicam que o valor esperado de uma intervenção em saúde resulta da soma de todos os valores esperados individuais que, por sua vez, advêm da ponderação do valor pela probabilidade de um determinado estado de saúde ocorrer. Como os indivíduos podem melhorar independentemente da intervenção ocorrer, fala-se de benefício líquido, i.e., da diferença de valor esperado com e sem intervenção em saúde que pode ser expressa da seguinte forma:

$$QALY = \sum_h \sum_t p_{1ht} Q_{ht} - \sum_h \sum_t p_{0ht} Q_{ht} \quad [2]$$

onde p_{0ht} representa a probabilidade do indivíduo se encontrar no estado de saúde h , no período t , sem ter sido sujeito a intervenção em saúde; p_{1ht} representa a probabilidade do indivíduo se encontrar no estado de saúde h , no período t , mas tendo sido sujeito a intervenção em saúde, e Q_{ht} representa o valor do estado de saúde h , no período t . Note-se que t assume a inclusão do desconto temporal, considerando

$$Q_{ht} = \frac{Q_h}{(1+r)^t} \quad [3]$$

e em que r corresponde à taxa de desconto temporal.

De acordo com Miyamoto (1999), uma aplicação típica do QALY envolve a análise de utilidade de uma decisão. Tome-se, como exemplo, um doente que escolhe uma de entre duas ou mais alternativas terapêuticas. Uma das componentes dessa análise constitui a construção de modelos de probabilidade para cada escolha de tratamento. Cada modelo descreve os possíveis estados de saúde que podem ocorrer associados à alternativa terapêutica e atribui uma probabilidade a cada um desses estados. Num caso simples, o modelo de probabilidade poderá ser descrito como uma lotaria para estados de saúde específicos. Em casos mais complexos, o modelo poderá ser apresentado como uma árvore de decisão ou um modelo de Markov. A segunda componente da análise de decisão envolve já a construção de um modelo de utilidade para os possíveis resultados em saúde. Juntando estas duas componentes de análise de decisão, i.e., considerando a elaboração de modelos de probabilidade para cada alternativa terapêutica e a concretização de um modelo de utilidade para os possíveis resultados em saúde, poder-se-á calcular a utilidade esperada de cada escolha de tratamento. O objectivo inerente à decisão será determinar qual das alternativas possui a utilidade esperada mais elevada (Miyamoto 1999). Neste âmbito, assume-se que o conceito central do QALY baseia-se nos princípios da teoria da utilidade, a teoria clássica de apoio à decisão em cenário de incerteza. A primeira formalização do modelo foi realizada por Pliskin e colaboradores (Pliskin et al. 1980), que identificaram as três condições para o QALY ser considerado uma medida cardinal válida de utilidade. O QALY é considerado uma função de utilidade válida, que representa as preferências individuais, se as seguintes condições forem garantidas (Brandeau et al. 2004):

- i) independência de utilidade mútua, i.e., a preferência por um de dois atributos, estando um deles fixo a um nível constante, não depende do nível particular desse atributo constante (e.g. um indivíduo com artrite reumatóide não julga o seu estado de saúde de forma diferente por ter cinco ou 20 anos de esperança de vida);
- ii) troca proporcional constante de anos de vida por estados de saúde, i.e., a proporção do tempo restante de vida que um indivíduo está disposto a trocar por

uma melhoria específica de qualidade é independente da quantidade actual de anos restantes de vida (e.g. se se considerar uma situação em que é pedido a um indivíduo para trocar uma certa quantidade do seu tempo de vida para ter um estado de saúde perfeita ao invés de um estado de saúde inferior, se ele disponibilizar 10 anos em 20, também disponibilizará cinco em 10 ou dois anos e meio em cinco);

iii) neutralidade face ao risco relativo à duração de vida, i.e., a função utilidade para os anos de vida é linear.

As condições identificadas por Pliskin e colaboradores podem ser simplificadas de acordo com as seguintes assumpções (Towers et al. 2005):

- Condição zero – os respondentes não têm preferências sobre os estados de saúde quando os anos de vida são igualados a zero, isto é, $(Estado\ de\ Saúde\ Q_1, 0) \sim (Estado\ de\ Saúde\ Q_2, 0)$;
- Escala intervalar – para permitir a agregação de QALY, os intervalos com igual distância têm a mesma interpretação, significando, por exemplo, que um ganho de 0,2 para 0,4 tem o mesmo valor que um ganho de 0,6 para 0,8;
- Neutralidade face ao risco – a preferência dos indivíduos não é afectada pelo nível de incerteza em torno do valor médio dos resultados.

Vários autores reconhecem que as condições acima elencadas são muito restritivas ou algo irrealistas, impossibilitando a obtenção das preferências dos indivíduos e resultando em decisões de tratamento sub-óptimas (Brazier et al. 2007; Brandeau et al. 2004; Towers et al. 2005). Veja-se, por exemplo, a condição da neutralidade face ao risco. Brandeau et al. (2004) assinalam alguns estudos que identificaram diferentes níveis de aversão ao risco nos seus sujeitos. Tome-se, como exemplos, o estudo de McNeil, Weichselbaum e Pauker (1978) que encontrou uma moderada aversão ao risco em relação aos anos de vida nos doentes com cancro pulmonar, o estudo de Stiggelbout et al. (1994)

que mostrou a existência de uma ligeira aversão ao risco em doentes com cancro do testículo e o trabalho de Verhoef, de Haan e van Daal (1994) que encontrou aversão ao risco em relação aos anos de vida em mulheres saudáveis mas preferências propensas ao risco em jogos envolvendo durações curtas de tempo. Já num contexto distinto, Mehrez e Gafni (1987) evidenciaram, por seu turno, a existência de aversão ao risco quando a duração de tempo foi maior. Nas situações apresentadas, a condição de neutralidade face ao risco não foi garantida podendo levar a que o QALY não fosse considerado uma representação válida das preferências dos indivíduos. No entanto, Brandeau et al. (2004) também apresentam as soluções propostas por outros investigadores que relaxam os critérios para considerar o QALY como uma função de utilidade válida. Esses investigadores sugerem, em primeiro lugar, que a neutralidade face ao risco e a condição zero são suficientes para garantir a existência e a validade de um modelo de QALY. A condição zero é, por si só, uma situação incontornável no contexto de saúde. Assim, subsiste a condição de neutralidade face ao risco para garantir a validação do modelo que, tal como referido anteriormente, é, por diversas vezes, incumprida. Miyamoto (1999) sugere que a condição zero e uma condição alternativa que designou como invariância de equivalentes de probabilidade (ou invariância de jogo-padrão) são suficientes para garantir o modelo multiplicativo do QALY. Dessa forma, um modelo generalizado de QALY manter-se-á válido, numa situação sem neutralidade face ao risco, se e só se ambas as condições de invariância de equivalentes de probabilidade e condição zero forem garantidas.

Quando se utilizam QALY, assume-se que o objectivo principal do decisor é maximizar a saúde da população, num contexto de recursos escassos. No entanto, Weinstein et al. (2009) referem que os decisores podem ter outros objectivos relacionados com equidade, justiça e metas políticas e para os quais o modelo convencional de QALY poderá não ser o mais adequado.

O QALY convencional expressa a utilidade pessoal dos resultados em saúde avaliados *ex ante* e em média pela população em geral que assume um véu de ignorância sobre a sua saúde futura. Como o valor expresso é obtido em termos de interesse pessoal e *ex ante*, i.e., por indivíduos que poderão não ter tido a experiência prévia com os estados

de saúde valorados, a decisão é habitualmente designada por decisão de utilidade (Nord et al. 2009).

No entanto, nada impede que os valores sejam elicitados *ex post*, i.e., por indivíduos com experiência directa corrente ou passada com os estados de saúde objecto de valoração (utilidade experienciada).

Além disso, o QALY pode ser modificado de forma a expressar o valor da sociedade quando estão em causa objectivos de justiça, ultrapassando assim o interesse pessoal (Nord et al. 2009).

Independentemente do modelo formal de análise escolhido, os resultados da avaliação económica não devem ser considerados como derradeiras respostas para a alocação de recursos, mas sim como contributos para futuras deliberações e processos de decisão justos.

2.2. Unidades de medida alternativas ao QALY

Ao longo dos últimos anos o conceito de QALY tem gerado alguma controvérsia e, por isso, outras unidades de medida têm sido sugeridas para introdução em técnicas de avaliação económica.

Apresentam-se, de seguida, algumas das unidades de medida alternativas aos QALY.

2.2.1. Anos saudáveis equivalentes

Os anos saudáveis equivalentes ou HYE (*Healthy Years Equivalent*) foram propostos no início da década de 1990, por Mehrez e Gafni, como uma melhoria teórica sobre os QALY (Mehrez & Gafni 1991). Tal como o QALY, esta unidade de medida combina dois resultados em saúde: qualidade de vida e longevidade.

Mehrez e Gafni (1991) defendem que os HYE representam na totalidade as preferências dos indivíduos pela forma como se calcula cada função individual de utilidade, ao não restringir a relação entre os estados de saúde e os anos de vida. São dois os pontos principais em que esta unidade de medida difere dos QALY. Em primeiro lugar, mede as

preferências dos indivíduos ao longo dos estados de saúde por que eventualmente passaria (caminho ou perfil) e não de cada estado de saúde individualmente; em segundo lugar, utiliza uma técnica de eliciação de preferências denominada de jogo-padrão em duas etapas, sendo que na primeira determina a utilidade convencional para o caminho de estados de saúde e na segunda mede o número de anos saudáveis que corresponde à mesma utilidade (Drummond et al. 2005).

Apesar do avanço teórico que representam, os HYE têm sido alvo de crítica. Um dos primeiros aspectos apontados relaciona-se com a viabilidade de realização da técnica do jogo-padrão por duas etapas. No entanto, tanto os autores que propõem esta unidade de medida, como outros, conseguiram replicar a técnica com relativa facilidade (Towers et al. 2005). A segunda crítica apontada, e mais relevante, é que a técnica de eliciação por duas etapas é redundante. Vários artigos de investigação referem que os HYE são teoricamente equivalentes à técnica de equivalência em tempo (Towers et al. 2005; Drummond et al. 2005). Por outro lado, a incorporação de risco na técnica de valoração utilizada (jogo-padrão) tem induzido alguma preocupação. Towers et al. (2005) referem ainda que os indivíduos poderão não expressar de forma rigorosa as suas preferências utilizando os HYE. Pelo facto de a tarefa ser cognitivamente exigente, os respondentes poderão ter alguma dificuldade em compreender a informação complexa contida num perfil de saúde e, dessa forma, recorrer a heurísticas para julgar esses perfis. Acresce ainda que as pessoas valoram alterações em saúde de acordo com o que pensam que vão sentir assim que essa alteração ocorre, ao invés de considerarem possíveis adaptações ao longo dos anos. Outra característica que se pode observar nos respondentes é a insensibilidade à abrangência do perfil, significando que poderão atribuir um valor inferior à totalidade do perfil do que à soma dos valores das partes que o constituem.

Com base na evidência científica que existe, não é claro que os HYE constituam uma melhor medida de preferências em saúde do que os QALY, justificando-se, assim, a realização de mais estudos de investigação nesta área (Towers et al. 2005).

2.2.2. Anos salvos equivalentes de vida jovem

Os anos salvos equivalentes de vida jovem ou SAVE representam uma alternativa ao QALY convencional. Esta unidade de medida foi introduzida tendo por base o argumento de que os QALY reflectem uma perspectiva individual e não uma perspectiva societal, deixando de parte o valor social (Drummond et al. 2005; Nord 1992).

A técnica utilizada no método do SAVE é a de equivalência em pessoas. Um exemplo da aplicação desta técnica passa, por exemplo, por questionar as pessoas sobre quantos doentes do tipo A deverão ser tratados para equivaler, em valor social, ao tratamento de 10 doentes do tipo B.

A abordagem dos SAVE difere da dos QALY na medida em que cada elemento da sociedade é questionado sobre que trocas gostaria para outros indivíduos, criando as bases da tomada de decisão societária. Nos QALY, cada elemento da sociedade é questionado sobre as trocas que gostaria para si mesmo e, nesse caso, a tomada de decisão societária toma forma a partir dessas trocas individuais (Drummond et al. 2005).

Não existe grande consenso teórico ou mesmo ético sobre qual das unidades de medida é mais vantajosa e vários investigadores defendem a necessidade de um aprofundamento da aplicação dos SAVE em mais estudos de investigação. Drummond et al. (2005) sugerem que, em alguns estudos, poderá ser uma possibilidade a introdução dos SAVE como medida complementar aos QALY, considerando que eventuais resultados distintos nos valores obtidos apenas enriquecerão a discussão dos agentes de decisão.

2.2.3. Anos de vida ajustados por incapacidade

Esta unidade de medida foi desenvolvida pela Organização Mundial de Saúde com o intuito de estimar o fardo global de doença e a efectividade das intervenções de saúde (Murray & Lopez 1996). Os anos de vida ajustados por incapacidade ou DALY (*Disability Adjusted Life-Year*) são conceptualmente semelhantes aos QALY, na medida em que combinam mortalidade e morbilidade numa só medida (Chen et al. 2015). O fardo de doença é medido em termos de DALY perdidos comparando com uma vida ideal saudável, com uma pontuação de qualidade equivalente a 1 e longevidade igual à do país com maior esperança de vida (no caso, o Japão). De forma a calcular os DALY totais para uma

população e determinada condição é necessário estimar os anos de vida perdidos, que resultam do número de mortes ocorridas numa população multiplicada pela média do remanescente da esperança de vida da população, aquando a idade da morte, e os anos vividos com incapacidade com severidade e duração conhecidas, associados a essa condição (Murray & Lopez 1996; Chen et al. 2015)

Apesar de conceptualmente semelhantes, os anos de vida ajustados por incapacidade apresentam algumas características que os distinguem dos QALY designadamente: a longevidade ser considerada constante (no QALY este valor difere em função da situação); os pesos associados à incapacidade basearem-se em pontuações da técnica de equivalência em pessoas obtidos através de um painel de profissionais de saúde (no QALY os pesos baseiam-se em preferências); os pesos poderem assumir, para além dos valores associados ao estado de saúde equivalente a morte e ao estado de saúde equivalente a saúde perfeita, apenas sete valores discretos (ao invés do QALY que pode assumir qualquer valor dentro da escala de utilidade) e utilizarem pesos de idade, isto é, ponderações que atribuem pesos inferiores a anos de vida de indivíduos jovens e idosos e ponderações que atribuem pesos superiores a anos de vida de indivíduos de meia-idade, pelo facto de serem estes que cuidam tanto de jovens como de idosos (Drummond et al. 2005).

O DALY como unidade de medida tem sido alvo de grande debate não só por poder gerar resultados com enviesamento positivo quando utilizado, por exemplo, em países com esperança de vida inferior à do país de referência (Japão), mas também pelas características acima descritas serem consideradas, por diversos investigadores, como arbitrárias e pouco transparentes (Drummond et al. 2005). Outra limitação significativa que influencia o uso de DALY é o método requerer uma diversidade significativa de informação relacionada com estruturas etárias específicas da população, esperança de vida, incidência, prevalência e outras métricas que dificulta a sua aplicação por parte dos investigadores (Chen et al. 2015).

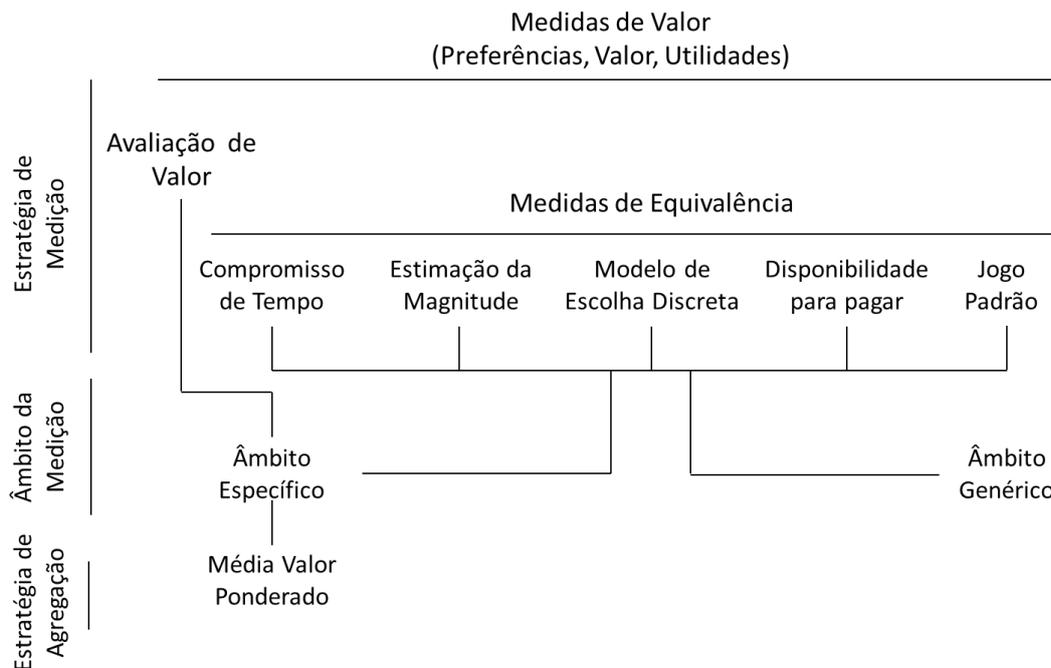
3. MEDIÇÃO DE PREFERÊNCIAS EM SAÚDE

O conceito de saúde, adoptado em 1946 pela Organização Mundial de Saúde, define a saúde de forma abrangente, considerando-a um estado de completo bem-estar físico, mental e social e não apenas a ausência de doença ou enfermidade (WHO 1946). Acresce ainda que, quando se está perante conceitos como taxas de sobrevivência e qualidade de vida relacionada com a saúde, as bases de decisão por parte dos indivíduos, profissionais de saúde e agentes de decisão ultrapassam os meros resultados biológicos e fisiológicos que uma intervenção em saúde pode proporcionar (Tsevat et al. 1994). Assim, a decisão sobre afectação de recursos em saúde pode ser dificultada se se tiver apenas em consideração os dados relativos ao impacto fisiológico de uma intervenção. O conceito abrangente de saúde pode efectivamente dificultar a decisão sobre o que deve conter um instrumento de medição em saúde (Brazier et al. 2007).

As técnicas para determinação da qualidade de vida relacionada com a saúde são distintas, podendo ser caracterizadas pela medição do estado de saúde do indivíduo (sistema descritivo que mede o estado de saúde) ou pela aferição de valor no que respeita a um determinado estado de saúde (Tsevat et al. 1994). Enquanto que a medição do estado de saúde mede objectivamente resultados fisiológicos ou inquire o indivíduo acerca da sua capacidade funcional e limitações, a medição de valor (valor, preferências, utilidade) questiona o indivíduo relativamente ao valor que ele atribui a um estado de saúde específico, assumindo que, por exemplo, indivíduos com limitações semelhantes poderão atribuir-lhes uma importância distinta. Drummond et al. (2005) explicam que apesar dos termos “valor”, “preferências” e “utilidade” serem, por vezes, utilizados de forma indistinta, há diferenças no seu significado. De acordo com os mesmos autores, o termo “preferência” aplica-se ao conceito mais lato, sendo que os termos “valor” e “utilidade” enquadram-se como tipos de preferências. A vantagem que Tsevat et al. (1994) apontam para a medição de valor face à medição do estado de saúde, é que a primeira parece ser o método que melhor se ajusta aos compromissos subjacentes ao processo de tomada de decisão.

Não obstante este capítulo debruçar-se apenas sobre as técnicas principais para valoração dos estados de saúde, na Figura 2 pode observar-se esquematicamente as técnicas disponíveis no processo para aferição do valor de um estado de saúde específico. A forma mais simples de aferir o valor de um estado de saúde é questionar o indivíduo para indicar numa determinada escala (e.g. de 0 a 100) a pontuação que lhe atribui. Alternativamente, pode questionar-se o indivíduo para comparar o estado de saúde com uma métrica explícita (tempo, dinheiro, outro estado de saúde ou disponibilidade para arriscar). Estas questões são denominadas medidas de equivalência.

Figura 2 - Aferição de Valor



Fonte: Adaptado de Tsevat et al. (1994). Using Health-related Quality-of-life Information: Clinical Encounters, Clinical Trials, and Health Policy. *Journal of General Internal Medicine*. 9(10), 576-582, Figura 1

Em 1944, von Neumann e Morgenstern, durante a sua investigação relativa à teoria dos jogos, desenvolveram uma teoria racional de tomada de decisão num contexto de incerteza, que ficou conhecida como a teoria da utilidade esperada ou teoria da utilidade de von Neumann-Morgenstern (von Neumann & Morgenstern 1944, 2007;

Drummond et al. 2005). Sob esta teoria, dada uma determinada alternativa (e.g. tratamento cirúrgico), estima-se um valor de utilidade para cada resultado possível, bom ou mau. Os valores de utilidade associados a cada resultado são multiplicados pela respectiva probabilidade de ocorrerem, sendo que a soma de todos eles representa a utilidade esperada da alternativa em análise (Brazier et al. 2007). O modelo normativo desenvolvido por estes investigadores determina o comportamento que um indivíduo, para ser racional, deve adoptar na tomada de decisões num contexto de incerteza sobre os resultados. Este modelo considera que os agentes de decisão possuem preferências que dominam um conjunto de possíveis escolhas, satisfazendo vários axiomas (Rulleau & Dachary-Bernard 2012; Lancsar & Louviere 2006; Drummond et al. 2005). Não obstante o debate que existe em torno destes axiomas e das alternativas que têm sido propostas, este modelo normativo persiste até aos dias de hoje como o modelo dominante, largamente implementado nos sectores público e privado (Drummond et al. 2005).

A teoria de utilidade esperada apresenta as condições sob as quais existe uma medida numérica de atractividade subjectiva de resultados (denominada de função de utilidade) com as seguintes propriedades: i) a função representa as preferências ordinais do decisor se os resultados forem obtidos num contexto de certeza; e ii) a ordem das utilidades esperadas associadas a várias decisões estratégicas num contexto de incerteza representam a ordenação dessas estratégias de acordo com as preferências do decisor (Brazier et al. 2007).

O Quadro 1 apresenta os axiomas definidos por von Neumann e Morgenstern expectáveis num comportamento racional em contexto de incerteza (Drummond et al. 2005; Lancsar & Louviere 2006).

Quadro 1 - Axiomas da teoria da utilidade de von Neumann-Morgenstern

Axioma	Definição
Integridade	O indivíduo tem bem definida a sua preferência entre duas possíveis lotarias (X e Y) ou é indiferente em relação a elas: $X \succ Y$ ou $X \sim Y$
Transitividade	Dadas três lotarias X, Y e Z, se $X \succ Y$ e $Y \succ Z$ então $X \succ Z$
Independência	Considerando a probabilidade $p \in (0,1)$ e qualquer lotaria X, Y e Z, $X \succ Y$ se e só se $pX + (1-p)Z \succ pY + (1-p)Z$. Esta condição significa que a preferência sobre duas lotarias não é afectada acrescentando uma terceira lotaria.

Axioma	Definição
Continuidade	Se $X \succcurlyeq Y \succcurlyeq Z$, então existe um $p \in (0,1)$ tal que $Y \sim pX + (1 - p)Z$. Esta condição pode ser complexa em situações extremas (e.g. quando X significa ganhar 105€, Y significa ganhar 100€ e Z significa morrer). O axioma da continuidade descarta as preferências lexicográficas e garante a existência de uma função utilidade.

De uma forma geral, o termo “utilidade” pode ser entendido numa lógica de quanto mais se preferir um determinado resultado, maior utilidade este terá associada. No entanto, este conceito revela-se complexo quando se tenta uma definição mais específica e, sobretudo, quando se procura a sua medição (Drummond et al. 2005)

A medição de preferências pode ser classificada em dois tipos: ordinal e cardinal. No tipo ordinal, os resultados são ordenados pelos indivíduos por ordem de preferência, de acordo com uma determinada graduação. No tipo cardinal, os indivíduos atribuem um número que representa a intensidade da sua preferência por um resultado e que se enquadra numa escala intervalar (Ferreira 2003; Drummond et al. 2005).

As técnicas utilizadas no processo de medição de preferências variam em função do tipo de preferências (valor ou utilidade) e do contexto de certeza/incerteza. Esquemáticamente, as técnicas para medição de preferências poderão ser representadas da seguinte forma (Quadro 2):

Quadro 2 - Técnicas de medição de preferências

Contextualização da Questão	
Cenário de Certeza (valores)	Cenário de Incerteza (utilidades)
Comparação em Pares	Jogo-Padrão
Equivalência em Tempo	
Equivalência em Pessoas	
Escala Categórica	
Escala de Pontuação	
Escala Visual Analógica	
Equivalência	
Escolha Discreta	
Estimação de Magnitude	

Fonte: Adaptado de Drummond et al. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 143, Tabela 6.1

As três técnicas principais para valoração dos estados de saúde são a escala visual analógica (*Visual Analogue Scale* ou VAS), o jogo-padrão (*Standard Gamble* ou SG) e a

equivalência em tempo (*Time Trade-Off* ou TTO). Embora mais recente, mas cada vez com maior utilização, tem-se também a experiência de escolha discreta (*Discrete Choice Experiment* ou DCE).⁷

Apresenta-se, de seguida, uma descrição de cada uma destas técnicas.

3.1. Visual Analogue Scale

A VAS é uma técnica proposta por Patrick e colaboradores para utilização em avaliação económica (Patrick et al. 1973). A utilização da VAS em avaliação económica teve início há cerca de três décadas, sendo uma das medidas mais utilizadas em vários instrumentos de medição em saúde, como o *Quality of Well-Being* (QWB), o *Health Utilities Index* (HUI), o 15-D ou o EQ-5D, com o intuito de valorar estados de saúde. Esta técnica consiste numa régua, normalmente desenhada na horizontal com 10 cm ou na vertical com 20 cm, com extremos bem definidos, utilizada com o objectivo dos respondentes indicarem um resultado em saúde. Os intervalos desta escala representam sempre a mesma diferença e correspondem à diferença em termos de preferência percebida pelos indivíduos (Brazier et al. 2007). A VAS pode ser utilizada para valorar estados crónicos considerados melhores que a morte, estados piores que morte e estados temporários. De forma a utilizar a VAS em avaliação económica é necessário garantir a comparabilidade entre respondentes. Para tal é necessário ter extremos claros e não ambíguos (saúde perfeita, morte). É também necessário definir o conceito de saúde perfeita para minimizar o risco de variação da interpretação entre respondentes e até mesmo dos investigadores (Brazier et al. 2007).

Para utilizar esta técnica é desejável que a valoração dos estados de saúde se situe numa escala entre 0 e 1, em que 0 representa estados equivalentes a morte e 1 representa o estado de saúde perfeita. No entanto, é também necessário permitir estados valorados

⁷ A maior parte da literatura refere-se à terminologia destas unidades de medida em inglês. Por esse motivo, ao longo do texto, serão utilizadas as respectivas siglas em inglês.

como piores que morte. Obtendo-se um valor para morte, então todos os estados de saúde são transformados de acordo o seguinte:

$$A_i = \frac{R_i - R_{morte}}{R_{saúde\ perfeita} - R_{morte}} \quad [4]$$

em que A_i corresponde à cotação da VAS ajustada para o estado de saúde h_i ; R_{morte} é a resposta atribuída ao estado de morte; R_i refere-se à resposta atribuída para o estado de saúde h_i e $R_{saúde\ perfeita}$ respeita ao melhor estado de saúde. O valor para A_i vai situar-se entre 0 e 1 ou poderá assumir valores negativos para estados de saúde considerados pior que morte (Brazier et al. 2007).

A determinação de preferências para estados de saúde temporários pode ser efectuada com recurso à VAS desde que explicado que os estados de saúde duram um período específico de tempo (semanas, meses ou anos desde que inferiores à esperança de vida) até se regressar a um estado saudável (Brazier et al. 2007)

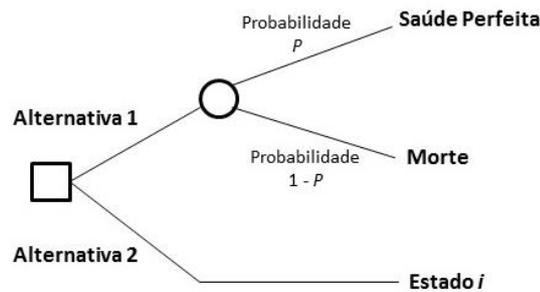
3.2. Standard Gamble

O SG constitui a técnica clássica de medição de preferências cardinais, baseando-se directamente nos axiomas da teoria de utilidade de von Neumann e Morgenstern (Torrance 1986). Na verdade, o jogo-padrão é uma aplicação directa do axioma da continuidade definido por estes investigadores (Drummond et al. 2005). Esta técnica tem sido utilizada de forma abrangente para apoio à decisão, particularmente no apoio à decisão médica dado o seu contexto de incerteza.

No SG o indivíduo é confrontado com duas alternativas: um resultado intermédio em saúde tido como garantido e um jogo com dois possíveis resultados (um melhor e o outro pior que o resultado intermédio). Esta técnica, utilizada para medição de estados de saúde temporários, estados de saúde crónicos, estados de saúde crónicos preferíveis à

morte ou piores que morte (o método sofre alterações ao seu formato em função da situação), pode ser ilustrada conforme apresentado na Figura 3.

Figura 3 - Jogo-padrão para um estado de saúde crónico preferível à morte

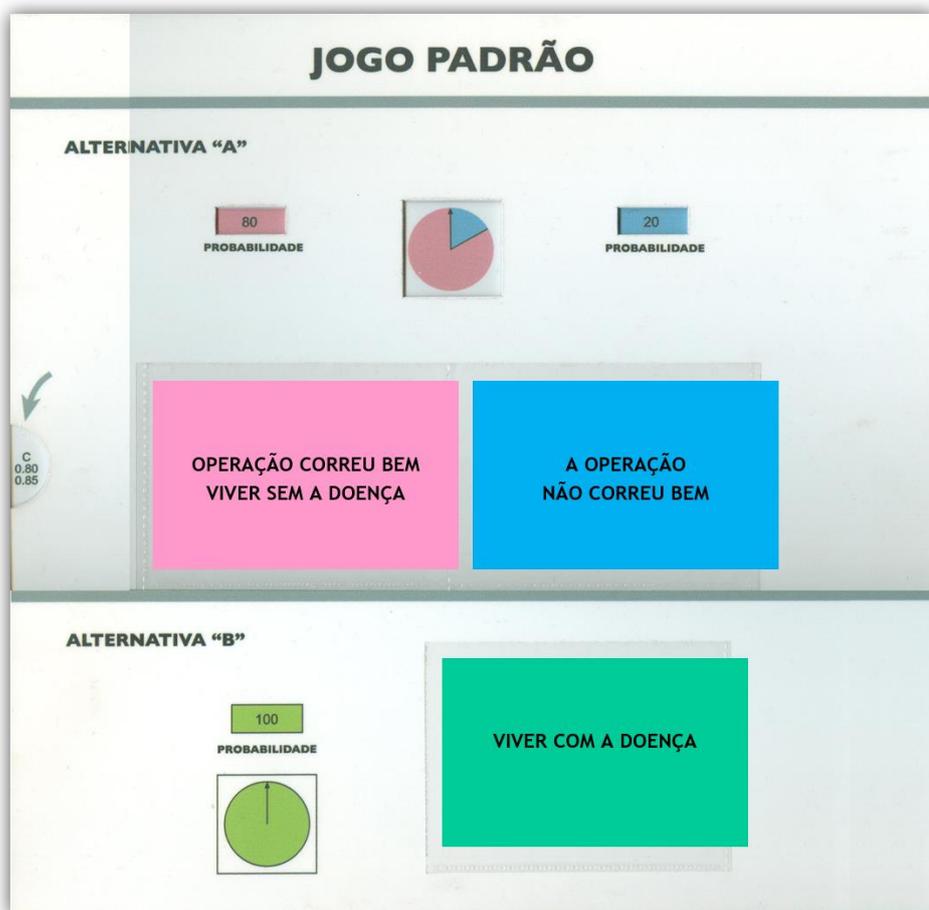


Fonte: Adaptado de Drummond et al. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 150, Figura 6.2

A alternativa 1 oferece dois resultados possíveis, com probabilidades distintas: ou o indivíduo retorna ao estado de saúde perfeita, vivendo t anos adicionais (probabilidade P) ou o indivíduo morre imediatamente (probabilidade $1 - P$). A alternativa 2 garante como resultado certo um determinado estado de saúde crónico i até ao final da vida (t anos). Neste jogo, a probabilidade P varia até o indivíduo ser indiferente entre as duas alternativas apresentadas. Neste caso, o valor do estado de saúde i (h_i) é precisamente igual ao valor de P medido numa escala de utilidade, em que a saúde perfeita para t anos corresponde a 1 e a morte imediata a 0 (Drummond et al. 2005; Ferreira 2003). Os algoritmos de iteração utilizados para a variação do valor de P até se alcançar a indiferença enquadram-se, habitualmente, nas categorias de trisseção (a duração de vida é sequencialmente alterada por incrementos ou decrementos fixos), bisseção (a duração de vida é sempre o ponto médio da secção remanescente da escala) e pingue-pongue (a duração de vida assume alternadamente valores altos e baixos).

Os resultados do jogo-padrão são obtidos através de entrevistas directas com os indivíduos, com recurso a ajudas visuais (e.g., quadros incluindo rodas de probabilidade), que ajudam na aferição das probabilidades (Figura 4).

Figura 4 - Quadro com roda de probabilidade para aplicação do SG



Nota: No jogo apresentado na figura, a Alternativa "A" corresponde a 20% de probabilidade de morrer na sala de operações e a Alternativa "B" significa permanecer com a doença o resto da vida.

As técnicas mais recentes já incluem o recurso a aplicações informáticas interactivas e a entrevistas de grupo com respostas obtidas em papel.

3.3. Time Trade-Off

A técnica TTO foi desenvolvida especificamente para utilização na área da saúde por Torrance e colaboradores (Torrance et al. 1972). Numa aplicação piloto efectuada por

estes investigadores, esta técnica originou valores semelhantes aos do SG, sendo considerada uma técnica mais simples de aplicar (Drummond et al. 2005; Torrance et al. 1972).

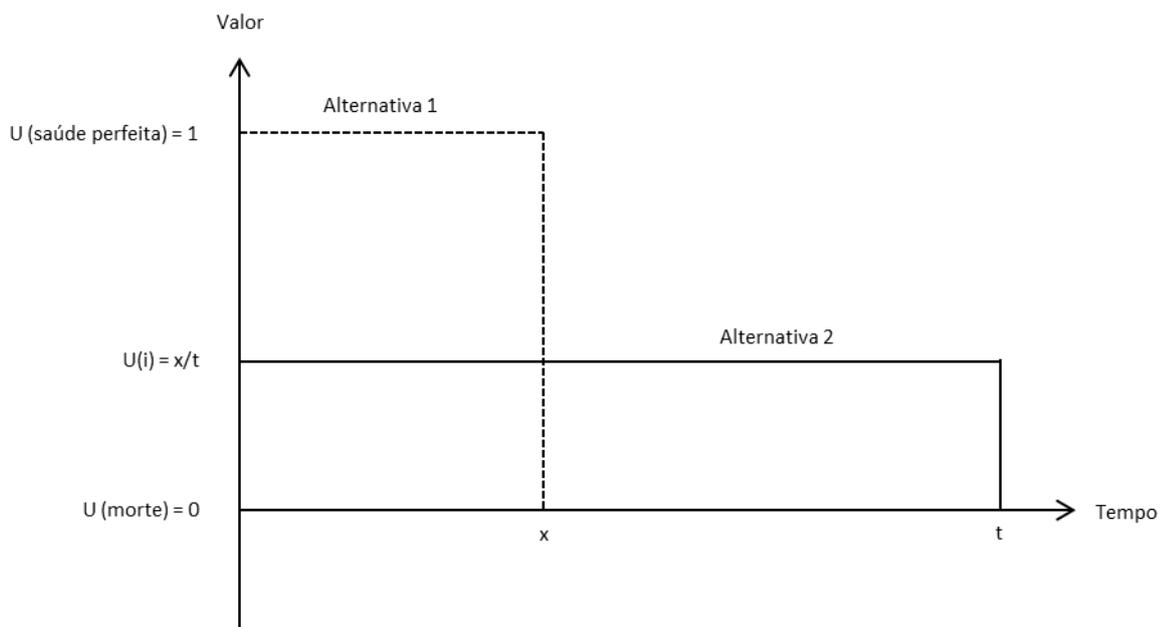
Estudos posteriores mostraram que, à luz da teoria de utilidade esperada de von Neumann e Morgenstern, para se considerarem como verdadeiras utilidades os resultados da técnica do TTO, alguns ajustes poderiam ter que ser efectuados aos valores brutos obtidos, assumindo como conhecido o tipo de função de utilidade dos indivíduos relativamente aos anos de vida (Johannesson et al. 1994; Martin & Lumley 2000; van Osch et al. 2004). Isto deve-se ao facto da função de utilidade de um indivíduo ser influenciada, entre outros factores, pela preferência temporal. Este conceito está presente quando o valor atribuído a um resultado é influenciado pelo momento do tempo em que esse resultado é obtido. De uma forma geral, as pessoas tendem a apresentar uma preferência temporal positiva, i.e., preferem resultados antes do que depois. Assumindo que o valor marginal por um período de tempo de sobrevivência adicional é constante, a preferência temporal pode fazer com que anos de vida adicionais pareçam apresentar um valor marginal decrescente na medida em que o valor presente desses anos futuros está reduzido pela preferência temporal positiva. A consequência deste efeito (desconto temporal) é modificar a função de utilidade da forma linear para côncava (Martin & Lumley 2000).

Enquanto que o SG pressupõe uma escolha entre um resultado certo e um jogo com base em incerteza, o TTO requer uma escolha mais simples entre duas alternativas tidas como certas. Trata-se de uma técnica menos complexa porque não obriga à compreensão de probabilidades por parte dos respondentes (Brazier et al. 2007). A equivalência em tempo é uma técnica que pode ser utilizada em estados de saúde crónicos considerados melhores ou piores que morte e em estados de saúde temporários (com as devidas adaptações à técnica).

Aos participantes de um exercício TTO é apresentada uma escolha entre duas alternativas. No caso de um exercício de valoração de um estado de saúde melhor que morte, com recurso ao TTO (Figura 5), a alternativa 1 corresponde a permanecer num estado de saúde perfeita por um período de tempo x (em que $x < t$), seguido de morte. A

alternativa 2 consiste em permanecer no estado de saúde i , pelo período de tempo t (e.g. 10 anos), ao qual se segue a morte. Nesta técnica, o tempo correspondente a x vai ser variado até que o indivíduo atinja o ponto de indiferença entre as duas alternativas oferecidas. Neste caso, o valor do estado de saúde i (U_i) será igual a $\frac{x}{t}$ (Ferreira 2003; Drummond et al. 2005; Oppe et al. 2016).

Figura 5 – TTO: estado de saúde melhor que morte

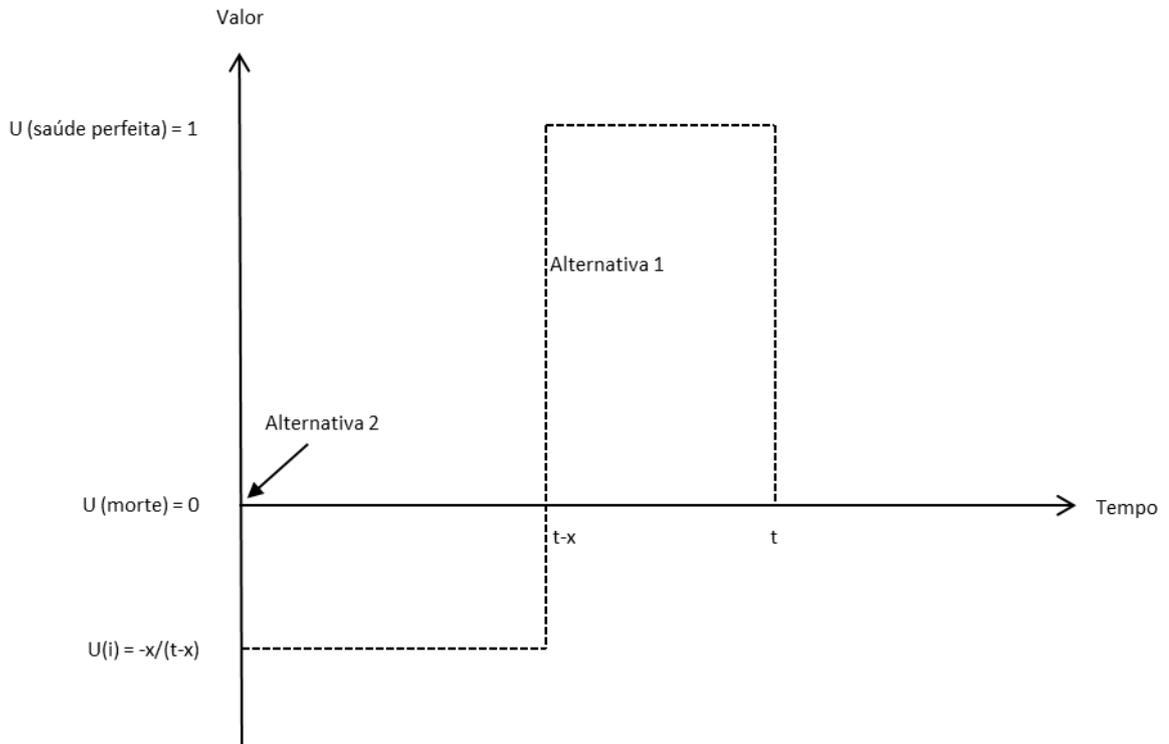


Fonte: Adaptado de Oppe et al. (2016). EuroQol protocols for time trade-off valuation of health outcomes. Pharmacoeconomics. (34) 993-1004, Figura 1a

Para a valoração de estados de saúde considerados piores que morte, é necessária uma abordagem ligeiramente diferente. O processo inicia-se com uma questão para determinar se um participante considera um determinado estado de saúde como sendo melhor ou pior que morte, pedindo ao participante para escolher entre morte imediata ou viver um determinado período de tempo num estado de saúde específico. Se o participante optar pela morte imediata, isso significa que considera o estado de saúde especificado como sendo pior que morte. Como à morte é habitualmente atribuída utilidade zero, ao estado de saúde apresentado deverá corresponder uma utilidade negativa (Attema et al. 2013). No TTO convencional, se o estado de saúde i é considerado pior que morte, o participante

fica perante duas alternativas (Figura 6): um cenário que envolve $(t - x)$ anos no estado de saúde i , seguido de x anos saudáveis e, por fim, a morte (alternativa 1) ou a morte imediata (alternativa 2). O valor de x é variado até ser encontrado o ponto de indiferença do participante que, neste caso, é dado por $-\frac{x}{(t-x)}$.

Figura 6 – TTO: estado de saúde pior que morte



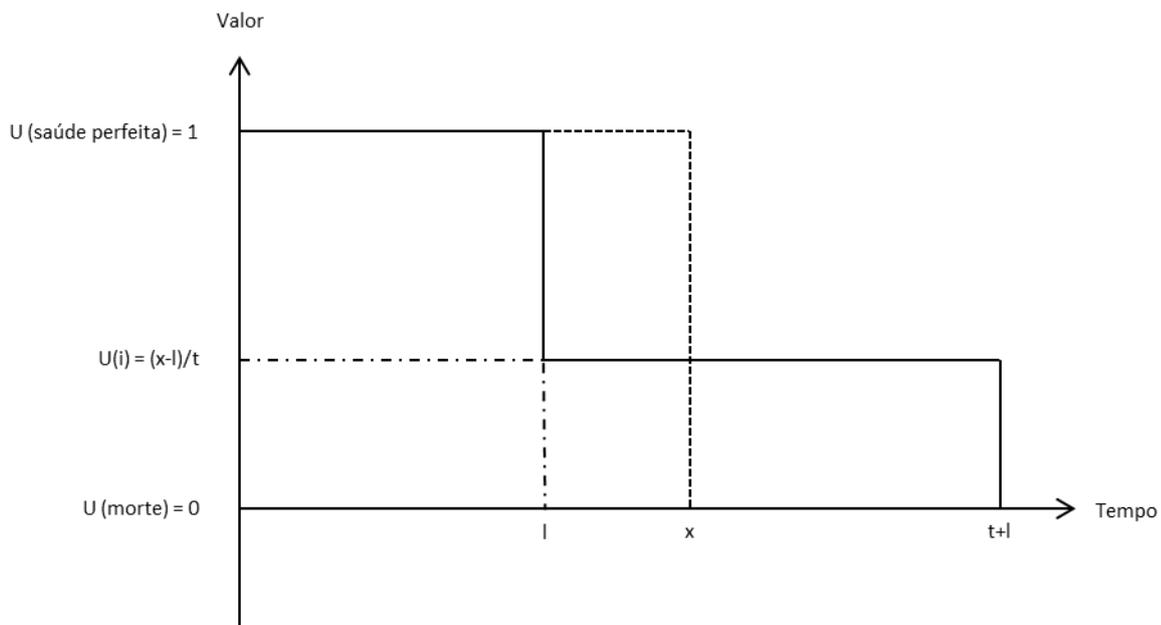
Fonte: Adaptado de Oppe et al. (2016). EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. (34) 993-1004, Figura 1b

Apesar da técnica do TTO ter uma aplicação abrangente, vários investigadores têm manifestado alguma preocupação pelo facto de se utilizarem processos diferentes para elicitação de preferências para os estados de saúde melhores que morte e estados de saúde considerados piores que morte (Augustovski et al. 2013; Oppe et al. 2016). Note-se que, para os primeiros, apenas é variado o tempo passado no estado de saúde perfeita enquanto que nos estados de saúde considerados piores que morte, a duração do tempo passado em saúde perfeita é variada simultaneamente com a duração do tempo passado no estado de saúde sob valoração. A questão levantada pelos investigadores é se se poderá

assumir que os resultados obtidos em ambos os processos integram a mesma escala, pelo facto de ambos os protocolos apresentados anteriormente para responder à valoração de estados de saúde considerados pior que morte poderem produzir valores extremamente negativos. Acresce ainda as regras arbitrárias que têm sido seguidas para converter esses valores extremamente negativos para uma escala cujo intervalo mínimo corresponde a -1.

De forma a resolver as questões elencadas, Robinson e Spencer apresentaram, em 2006, uma nova forma de elicitación capaz de produzir simultaneamente valores superiores e inferiores a zero (Robinson & Spencer 2006). O processo consistia em adicionar tempo adicional em saúde perfeita (l) ao estado de saúde de vida perfeita (Alternativa 1) e também ao estado de saúde a ser valorado (Alternativa 2 ou h_i). Neste último estado, o tempo em saúde perfeita precede o estado de saúde a ser valorado, sendo este método designado como *lead-time* TTO. Quando o estado de saúde valorado é considerado melhor que morte, o ponto de indiferença é alcançado quando a duração do tempo de vida na Alternativa 1 é superior ao número de anos em saúde perfeita oferecido na Alternativa 2 (Figura 7).

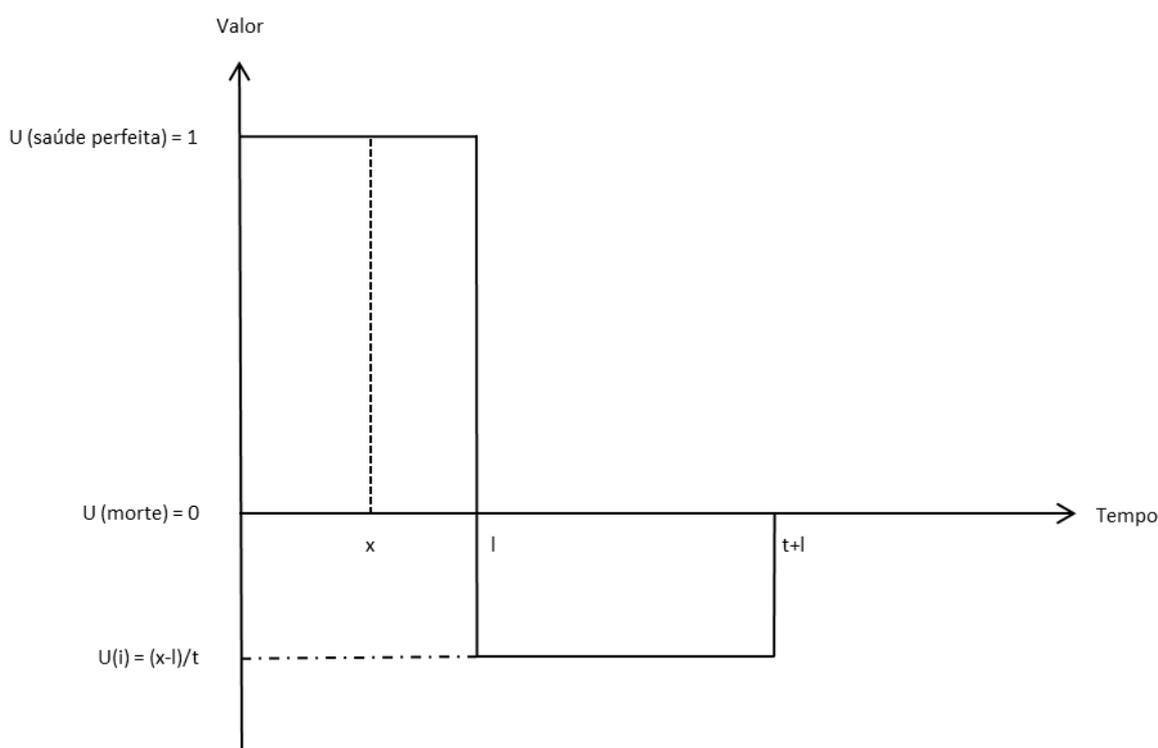
Figura 7 - *Lead-time* TTO: estado de saúde melhor que morte



Fonte: Adaptado de Oppe et al. (2016). EuroQol protocols for time trade-off valuation of health outcomes. Pharmacoeconomics. (34) 993-1004, Figura 2a

Se o estado for considerado pior que morte, o ponto de indiferença situar-se-á quando a duração do tempo de vida na Alternativa 1 for inferior ao número de anos em saúde perfeita apresentado na Alternativa 2 (Figura 8). Ao contrário da abordagem convencional do TTO, este processo iterativo permite ao participante transitar entre valores negativos e positivos sem ter que explicitamente pensar se o estado de saúde é pior ou melhor que morte (Robinson & Spencer 2006).

Figura 8 - *Lead-time* TTO: estado de saúde pior que morte



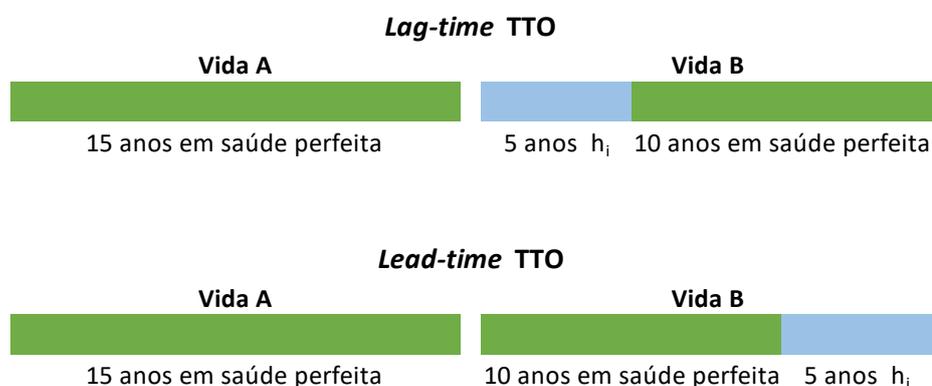
Fonte: Adaptado de Oppe et al. (2016). EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. (34) 993-1004, Figura 2b

Tilling e colaboradores apresentaram ainda uma variante denominada de *lag-time* TTO que, ao contrário do processo anterior, apresenta o tempo adicional em saúde perfeita cronologicamente depois do estado de saúde a ser valorado (Tilling et al. 2010).

Alguns autores sugerem que, em teoria, os valores obtidos com recurso ao *lead-time* e ao *lag-time* TTO deverão ser semelhantes, no entanto, recomendam a realização de mais estudos para confirmação desta assumpção (Augustovski et al. 2013).

Na Figura 9 é possível observar as duas variantes – *lead-time* e *lag-time* – da técnica do TTO.

Figura 9 - *Lead-time* e *Lag-time* TTO



Fonte: Adaptado de Augustovski et al. (2013). Lead versus lag-time trade-off variants: Does it make any difference? *European Journal of Health Economics*. 14 (Sup. 1), S25-S31, Figura 1

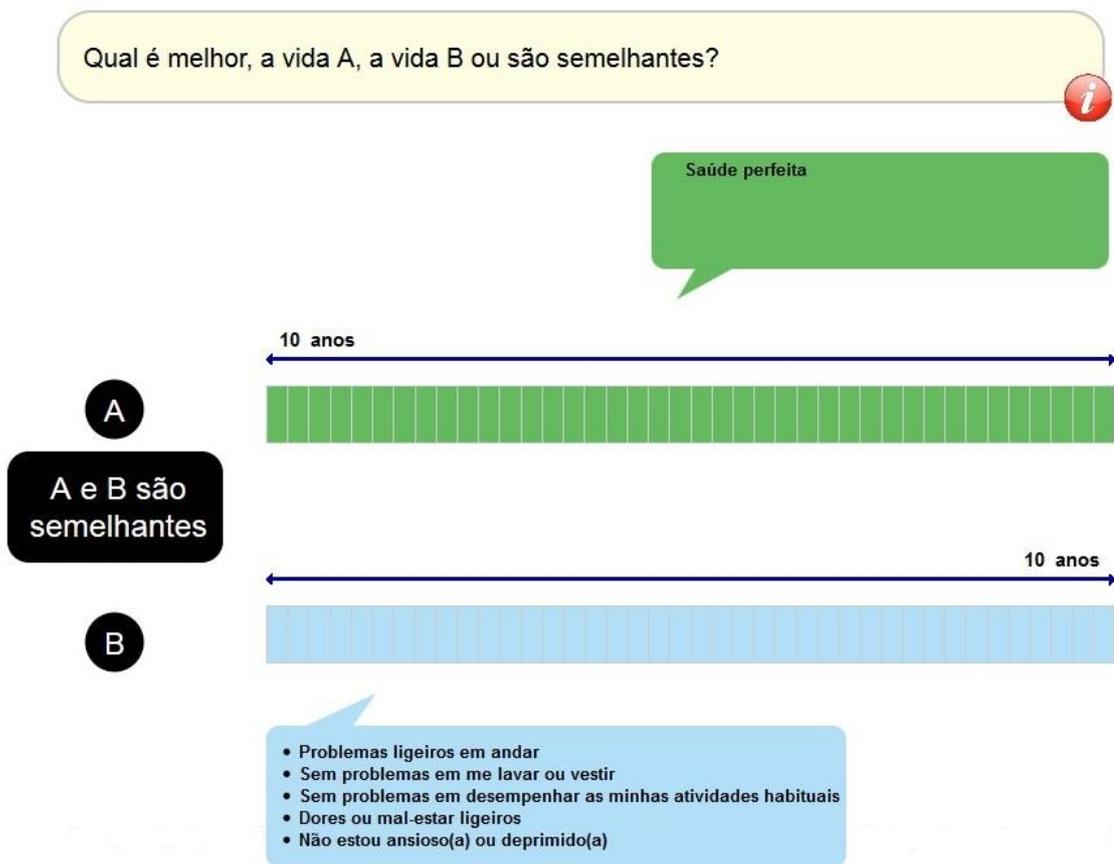
Estas variantes não são, no entanto, livres de críticas nomeadamente porque sofrem de efeitos de enquadramento, i.e., quanto maior o *lead-time* oferecido, mais tempo o participante troca. Os participantes podem ainda utilizar todo o tempo da escala para troca mesmo em situações cujos estados de saúde não são muito severos. Adicionalmente, alguns participantes, provavelmente por dificuldade na compreensão das variantes, optam pelo caminho mais fácil e encurtam a tarefa passando apenas por um ou dois passos durante o processo de movimentos iterativos e outros poderão não distinguir entre estados de saúde melhores e piores que morte quando trocam *lead-time* (Augustovski et al. 2013; Oppe et al. 2016).

Na sequência dos problemas detectados em vários estudos de investigação, na aplicação destas variantes, o grupo EuroQol desenvolveu uma nova abordagem designada de TTO compósito (cTTO) que combina o protocolo convencional do TTO para obter valores superiores a zero e a variante *lead-time* para os valores inferiores a zero, conforme apresentado nas Figuras 10 e 11 respectivamente (B. Janssen et al. 2013).

A variante *lead-time* TTO foi preferida em detrimento da *lag-time* por considerar-se que, conceptualmente e em termos práticos, a primeira encontra-se mais alinhada com

o TTO convencional e por ambas as variantes (*lead* e *lag-time*) terem resultados semelhantes. Nesta abordagem, a distinção entre estados de saúde melhores e piores que morte é efectuada através de duas tarefas TTO distintas: TTO convencional para eliciar valores melhores que morte e *lead-time* TTO para eliciar valores piores que morte. O cTTO pressupõe que todas as tarefas de valoração se iniciam da mesma forma, i.e., com recurso ao TTO convencional: 10 anos no estado que está a ser valorado (Vida B) e 10 anos em saúde perfeita (Vida A).

Figura 10 - cTTO. a) TTO convencional para valoração de estados de saúde melhores que morte

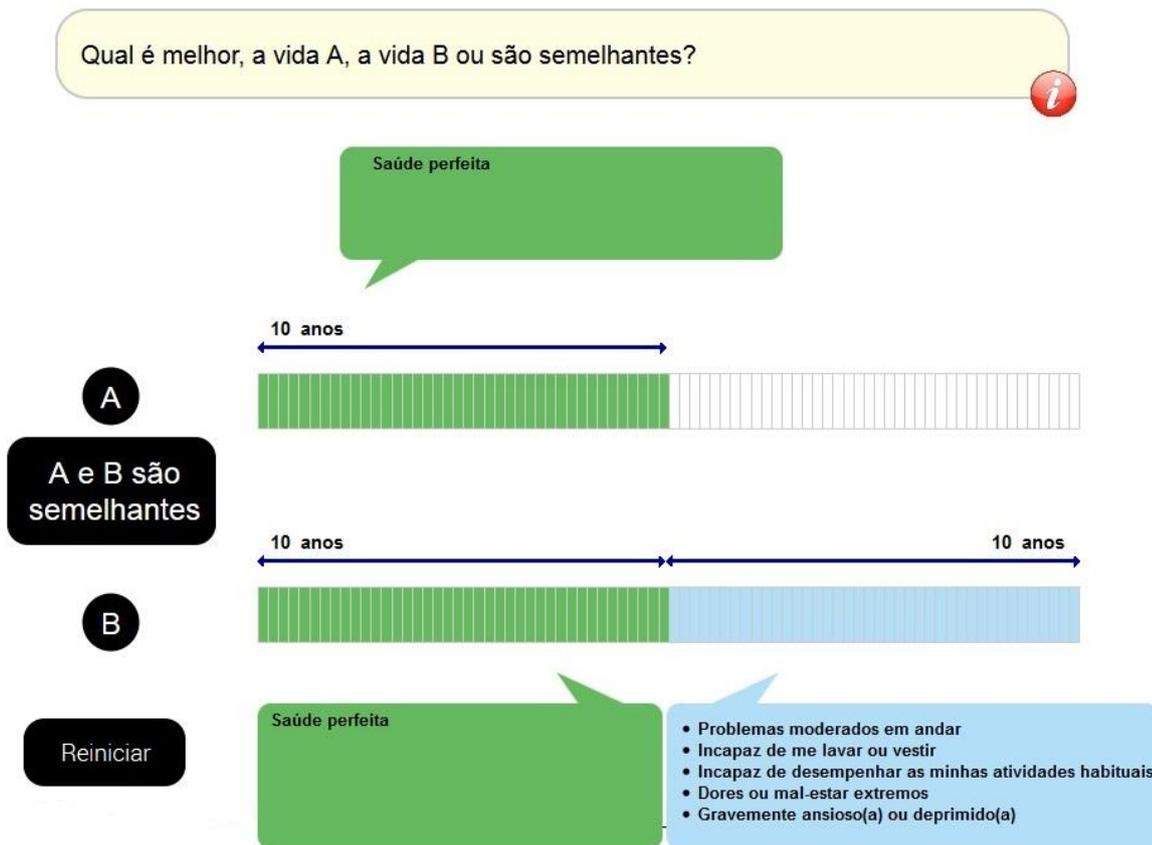


Fonte: Retirado da versão portuguesa do software EQ-VT (EuroQol Group Valuation Technology)

Para estados de saúde piores, os participantes podem escolher trocar a totalidade dos 10 anos em saúde perfeita, assumindo-se que o estado a ser valorado é pelo menos igual a zero. Nestas circunstâncias e com o intuito de eliciar valores inferiores a zero, a variante *lead-time* TTO é introduzida, adicionando mais 10 anos ao estado de saúde

perfeita e, correspondentemente, mais 10 anos em saúde perfeita antes de se alcançar o estado de saúde a ser valorado (B. Janssen et al. 2013).

Figura 11 - cTTO. b) *Lead-time* TTO para valoração de estados de saúde piores que morte



Fonte: Retirado da versão portuguesa do software EQ-VT (EuroQol Group Valuation Technology)

Note-se que, em geral, a abordagem TTO assume a condição de troca proporcional constante de anos de vida, implicando que a quantidade de tempo trocada é independente da duração temporal considerada. O cTTO assume, para além desta, a condição adicional de independência aditiva, que implica que o valor associado a um estado de saúde no período T_1 é independente do valor de um outro estado de saúde no período T_2 (B. Janssen et al. 2013). Assim, os valores brutos obtidos para o cTTO são transformados assumindo-se iguais ao tempo passado no estado de saúde acrescentado de 10 anos (em avanço ou *lead-time*) com saúde perfeita e transformados de acordo com a fórmula seguinte:

$$Valor\ cTTO = \frac{(valor\ bruto - 10)}{10}$$

[5]

A escolha de 10 anos como horizonte temporal para a duração de tempo no estado de saúde a ser valorado foi arbitrariamente definida pelo grupo EuroQol. No entanto, esta escolha requer uma análise e investigação aprofundadas na medida em que o horizonte temporal pode influenciar os resultados obtidos pela técnica de valoração, o que conflua com a condição de troca proporcional constante de anos de vida. Oppe e colaboradores (2016) defendem que poderão ser testados distintos horizontes temporais, fixos e variáveis. Adicionalmente, sugerem intervalos de tempo mais curtos para determinados estados de saúde em que se considere plausível que os indivíduos permanecem durante um curto período de tempo neles. Por outro lado, um horizonte temporal de 10 anos pode ser irrealista para indivíduos idosos ou com patologias graves, situações que podem sugerir uma esperança de vida mais curta que o horizonte temporal definido. Além disso, a escolha de intervalos de tempo mais curtos pode conduzir à impossibilidade de equivaler os resultados com os actuais obtidos com 10 anos de horizonte temporal. Todas estas questões obrigam a uma reflexão profunda e a investigações futuras.

A aplicação do cTTO demonstrou ser uma abordagem mais consistente na elicitação de valores negativos, sem necessidade de recorrer a um redimensionamento discricionário da escala de valores, tal como sucedia com o TTO convencional ao permitir uma escolha arbitrária de tempo para troca (B. Janssen et al. 2013). No entanto, mesmo esta abordagem tem as suas limitações. Em primeiro lugar, por assumir a condição extra de independência aditiva, para além da condição de troca proporcional constante. Em segundo lugar, por consistir em dois modos diferentes de elicitação de valores: uma para estados de saúde melhores que morte e outra para estados de saúde piores que morte. No entanto, poder-se-á sempre considerar que, para qualquer estado de saúde em valoração, o participante tem sempre a possibilidade de introduzir a variante *lead-time* TTO utilizando, assim, um só modo de elicitação (B. Janssen et al. 2013).

De acordo com Oppe et al. (2016), a variação da duração do tempo de vida na Alternativa 1 (saúde perfeita) carece de três requisitos: determinação de um ponto de

partida, a existência de um algoritmo de iteração e uma regra para a sua terminação. O ponto de partida pode assumir um valor fixo ou aleatório enquanto que o algoritmo de iteração e a regra de terminação são normalmente estandardizados para todos os participantes. Os algoritmos de iteração utilizados inserem-se nas categorias de trisseccção, bissecção e ping-pong, assumindo-se combinações destes na técnica de valoração. A regra para a sua terminação pode ser a identificação do ponto de indiferença ou os limites de um intervalo em torno do ponto de indiferença (Oppe et al. 2016).

Para concretizar a valoração de estados de saúde, o grupo EuroQol desenvolveu três protocolos que incluem o TTO como técnica de valoração principal: i) o protocolo *Measurement and Valuation of Health* (MVH); ii) o protocolo de Paris e iii) o protocolo *EuroQol Valuation Technology* (EQ-VT). O Quadro 3 sumariza as principais características dos três protocolos.

Quadro 3 - Principais características dos protocolos MVH, Paris e EQ-VT

	MVH	Paris	EQ-VT
Estrutura TTO:			
Estrutura	TTO convencional		cTTO
Estados de Âncora	11111 (ausência de problemas nas dimensões do EQ-5D-3L)		“Saúde perfeita” e “Morte”
Horizonte temporal	10 anos para estados de saúde melhores e piores que morte		10 anos para estados de saúde melhor que morte e 20 anos para estado de saúde piores que morte
Estrutura das vidas hipotéticas	Estados de saúde melhores que morte: Vida A – viver em saúde perfeita durante x anos seguidos de morte (Alternativa 1); Vida B – viver num estado de saúde limitado durante 10 anos seguidos de morte (Alternativa 2) Estados de saúde piores que morte: Vida A – viver num estado limitado durante $10 - x$ anos, seguidos de saúde perfeita durante x anos, seguidos de morte (Alternativa 1); morte imediata (Alternativa 2)	Estados de saúde melhores que morte: Vida A – viver em saúde perfeita durante x anos seguidos de morte (Alternativa 1); Vida B – viver num estado de saúde limitado durante 10 anos seguidos de morte (Alternativa 2) Estados de saúde piores que morte: Vida A – viver em saúde perfeita durante x anos, seguidos de morte (Alternativa 1); Vida B – viver em saúde perfeita durante 10 anos, seguidos do estado limitado durante 10 anos, seguidos de morte (Alternativa 2)	

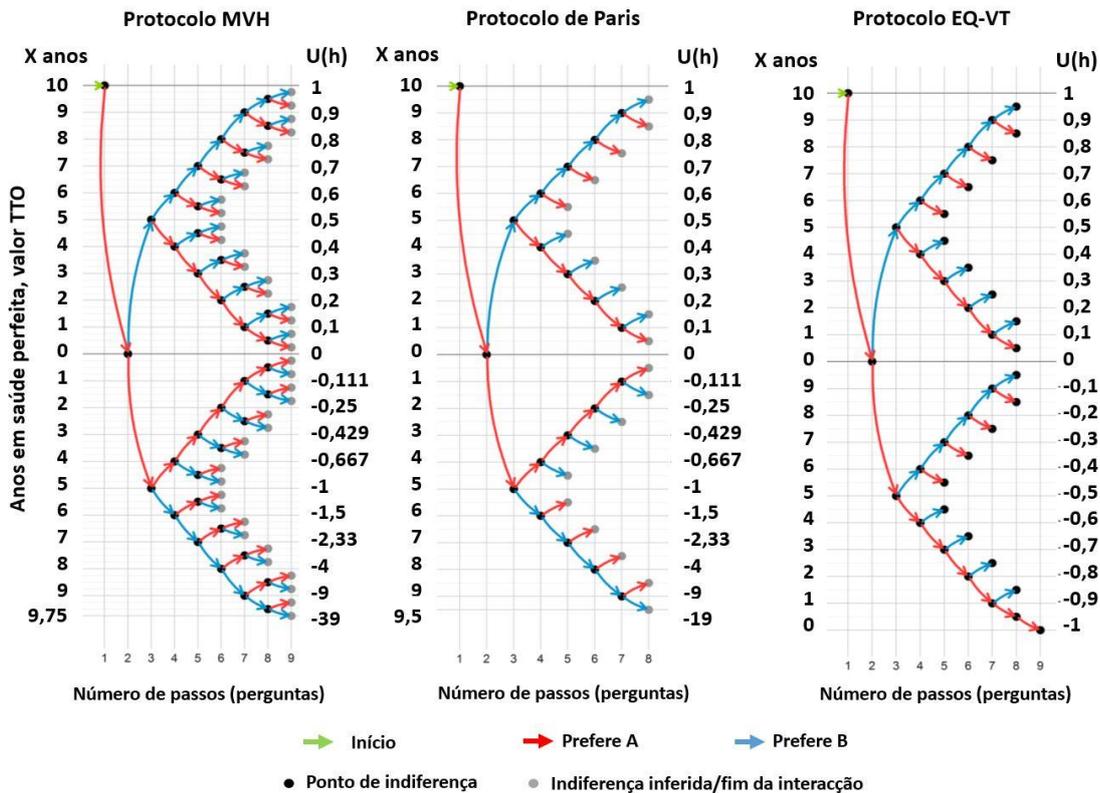
	MVH	Paris	EQ-VT
Procedimentos de iteração, modo e método de administração:			
Ponto de partida da iteração	10 anos		
Algoritmo de iteração	Bisseção para os primeiros três passos seguida de trisseção crescente/decrecente com 1 ano de incremento e uma correcção de 6 meses quando ocorre a reversão de preferências	Bisseção para os primeiros três passos seguida de trisseção crescente/decrecente com 1 ano de incremento, sem correcção	Bisseção para os primeiros três passos seguida de trisseção crescente/decrecente com 1 ano ou 6 meses de incrementos e uma correcção de 6 meses quando ocorre a reversão de preferências
Regra de terminação da iteração	Indiferença/intervalo de indiferença = 6 meses	Indiferença/intervalo de indiferença = 1 ano	Indiferença
Modo de administração	Entrevistas presenciais e individuais		Entrevistas presenciais e individuais assistidas por computador
Método de administração e ajudas visuais	Lápis e papel e um quadro temporal como ajuda visual		Recolha de dados electrónicos e apresentação digital como ajuda visual
Treino e controlo de qualidade:			
Treino do participante	Tarefas de aquecimento: classificação e VAS seguidas da explicação TTO		Um exemplo de treino e três exemplos práticos
Treino do entrevistador	Aula prática não estandardizada		
Controlo de qualidade	Não foram formalizados nem incluídos procedimentos de controlo de qualidade		<i>Software</i> estandardizado de controlo de qualidade

Fonte: Adaptado de Oppe et al. (2016). EuroQol Protocols for Time Trade-Off Valuation of Health Outcomes. *PharmacoEconomics*. 34 (10), 993-1004, Tabela 1

Os protocolos MVH e de Paris, criados respectivamente em 1993 e 2009, foram desenvolvidos com o intuito de valorar os estados de saúde originados pelo instrumento de medição EQ-5D-3L com recurso ao TTO convencional. O protocolo EQ-VT surgiu em 2012 para valoração dos estados de saúde decorrentes do instrumento de medição em saúde EQ-5D-5L. Não obstante o grupo EuroQol considerar existir potencial de melhoria para o protocolo EQ-VT, este é actualmente assumido como sendo o estado da arte (Oppe et al. 2016).

A Figura 12 apresenta os esquemas de iteração associados aos três protocolos de valoração do grupo EuroQol.

Figura 12 - Esquemas de iteração dos protocolos de valoração do EuroQol



Fonte: Adaptado de Oppe et al. (2016). EuroQol Protocols for Time Trade-Off Valuation of Health Outcomes. *PharmacoEconomics*. 34 (10), 993-1004, Figura 5

Na figura, h corresponde ao estado de saúde a avaliar, $U(h)$ é o valor do estado de saúde h e x respeita ao tempo em saúde perfeita (anos). Tal como Oppe e colaboradores (2016) explicam, os três protocolos apresentam, para estados de saúde considerados melhores que morte, a vida A durante x anos em saúde perfeita e a vida B durante 10 anos no estado de saúde em valoração. Para estados de saúde considerados piores que morte, os protocolos MVH e de Paris apresentam a vida A como representando $10 - x$ anos no estado de saúde em valoração seguidos de x anos em saúde perfeita e a vida B representando a morte imediata. Nos estados de saúde piores que morte, a diferença para o protocolo EQ-VT reside no facto da vida A ser apresentada como x anos em saúde perfeita e a vida B como 10 anos em saúde perfeita seguidos de 10 anos no estado de saúde em valoração. Nos protocolos MVH e de Paris, as tarefas de valoração seguem ao longo da direcção das setas até a indiferença ou o fim do ramo serem alcançados. No protocolo EQ-

VT, a tarefa de valoração termina apenas quando a indiferença é indicada pelo participante, sendo que a direcção da tarefa pode ser revertida a qualquer momento.

3.4. Limitações das técnicas de medição de preferências mais comuns

Tal como referido anteriormente, a VAS, o SG e o TTO consistem nas técnicas mais utilizadas para eliciação de preferências em saúde. No entanto, tem surgido um intenso debate sobre qual será a melhor técnica. Ainda que o TTO e o SG sejam técnicas reconhecidamente superiores à VAS, pelo facto de se basearem em escolhas que envolvem sacrifício ou uma troca, os valores elicitados por cada uma destas técnicas geram pouco consenso (Rowen et al. 2015). Van Osch e colaboradores (Van Osch et al. 2004) referem que os valores de utilidade obtidos por estas técnicas podem ser enviesados por questões relacionadas com aversão ao risco, compatibilidade de escala, curva de utilidade e peso das probabilidades. Estas preocupações foram também partilhadas por outros autores (Brazier et al. 2012). O Quadro 4 resume o tipo de enviesamento e o respectivo efeito associados às técnicas SG e TTO:

Quadro 4 - Tipo de enviesamento e efeito por técnica de eliciação de preferências

Técnica	Curva de Utilidade	Peso da Probabilidade	Aversão ao Risco	Compatibilidade de Escala	Efeito Total
TTO	Convexa	Não aplicável	Côncava	Côncava	?
SG	Não aplicável	Côncava (quase sempre)	Côncava	Desconhecido	Côncava

Fonte: Adaptado de van Osch et al. (2004). Correcting Biases in Standard Gamble and Time Tradeoff Utilities. *Medical Decision Making*. 24(5), 511-517, Tabela 1

No que concerne à curva de utilidade, o TTO assume uma função linear da duração de vida, o que nem sempre é verdade na medida em que grande parte dos indivíduos atribui maior valor aos anos próximos do que aos anos futuros (função utilidade côncava e não linear). Como na técnica do TTO os indivíduos são questionados sobre a troca de anos futuros, estes anos serão, conseqüentemente, sobrevalorizados nos cálculos, resultando num enviesamento negativo das utilidades.

O peso da probabilidade apenas respeita ao SG. Neste tipo de enviesamento, observa-se que os indivíduos tendem a sobrevalorizar probabilidades baixas (inferiores a 0,33) e a subvalorizar probabilidades mais elevadas (superiores a 0,33). Dado que as utilidades associadas aos estados de saúde excedem, habitualmente, os 0,33, este enviesamento conduz a uma sobrestimação dos valores de utilidade obtidos através do SG.

Quanto à aversão ao risco, ambas as técnicas (TTO e SG) conduzem a enviesamentos positivos, sobrestimando valores de utilidade associados aos estados de saúde. No caso do TTO, um indivíduo com aversão ao risco estará mais relutante em trocar anos de vida (considerado como perda) por saúde óptima (ganho). No caso do SG e considerando ainda um indivíduo com características de aversão ao risco, o ganho associado à probabilidade P terá que ser muito elevado para compensar a perda associada à probabilidade $1 - P$, gerando assim um enviesamento positivo nas probabilidades.

A compatibilidade de escala respeita à constatação de que quanto maior for a compatibilidade de uma característica com a escala de resposta utilizada, maior atenção e maior peso serão atribuídos pelo indivíduo a essa característica. No que concerne ao TTO, a escala de resposta refere-se ao número de anos em saúde perfeita. Neste caso, uma atenção particular será dada à duração do estado de saúde, ao invés do impacto do estado na saúde do indivíduo, resultando em valores mais elevados. No caso do SG, a escala de resposta é a probabilidade. No entanto, o comportamento e respectivo efeito no enviesamento são muito difíceis de estimar (Van Osch et al. 2004).

De uma forma geral, desconhece-se o efeito cumulativo destas limitações na obtenção de valores de utilidade através do TTO e assume-se que os valores de utilidade obtidos através do SG são mais elevados. Não obstante alguns métodos terem sido propostos com vista à correcção dos enviesamentos apresentados, nem todos os efeitos foram corrigidos (Van Osch et al. 2004).

O TTO e o SG têm sido também criticados pela sua complexidade na aplicação aos respondentes e por excluírem grupos particulares como crianças, idosos ou culturas diferentes (Brazier et al. 2012; Rowen et al. 2015). Por outro lado, a VAS é uma técnica que nem sequer é estritamente baseada em preferências na medida em que não implica uma escolha entre opções (Brazier et al. 2007).

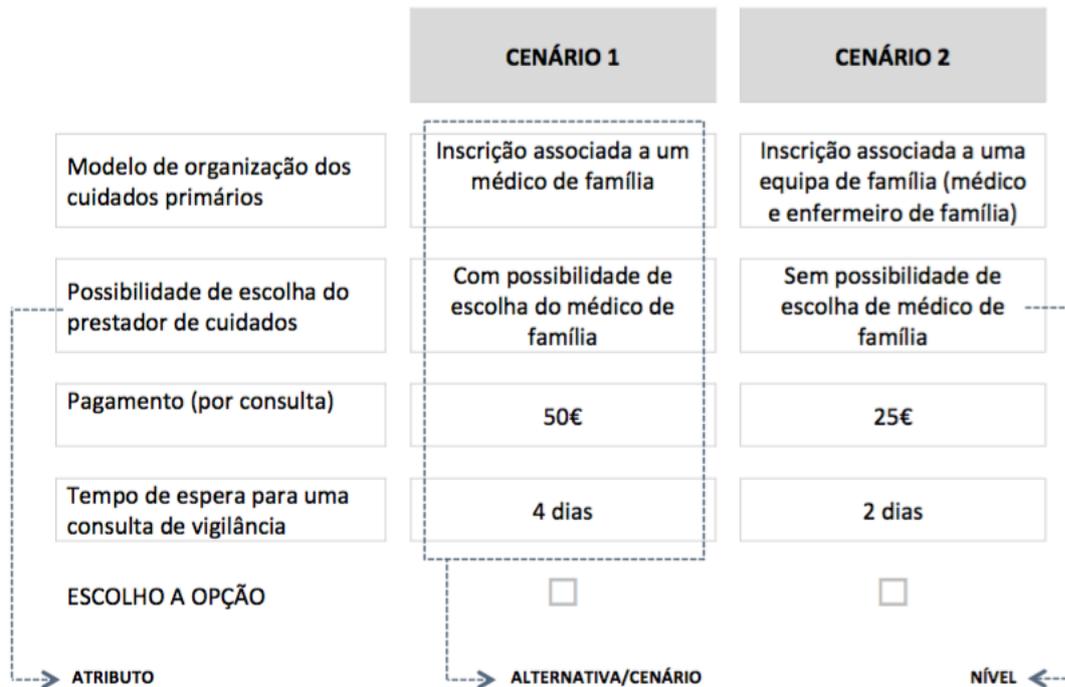
As limitações destas técnicas têm conduzido a um interesse crescente por outras tais como a Estimação de Magnitude, Equivalência em Pessoas e DCE. Apesar destas técnicas serem menos frequentemente utilizadas, a DCE tem vindo a obter cada vez mais destaque em estudos para eliciação de preferências em saúde.

3.5. Discrete Choice Experiment

A DCE é uma técnica utilizada desde o início da década de 60 do século passado, que assenta na psicologia matemática e cuja aplicação iniciou-se originalmente na área do marketing. Desde essa altura, a DCE tem sido utilizada para eliciar as preferências dos indivíduos em outras áreas académicas como a economia dos transportes, economia do ambiente e, mais recentemente, a economia da saúde (Earl & Kemp 1999; Kenny et al. 2003; Lancsar & Louviere 2008; Norman, Viney, et al. 2013). A utilização desta técnica de eliciação é particularmente relevante pelo facto de as preferências dos indivíduos poderem constituir um papel fundamental para a determinação das prioridades na aplicação dos recursos em saúde (Lancsar & Louviere 2008). Adicionalmente, a utilização da DCE justifica-se pela sua simplicidade de aplicação e pela compreensão facilitada do método (Kjær 2005).

A escolha discreta ocorre quando um indivíduo se encontra face a uma opção entre um conjunto de alternativas finitas, mutuamente exclusivas e que contemplam todas as possibilidades (Antunes et al. 2017). A DCE é habitualmente implementada com recurso a questionários que definem o bem ou serviço em análise em função de atributos, considerados como os mais relevantes, aos quais o indivíduo associa uma utilidade. O objectivo desta técnica é observar a resposta do indivíduo a alterações nesses atributos através de níveis que lhes são atribuídos. Cada combinação, construída com os vários atributos e níveis associados, vai constituir um cenário cujo objectivo final é permitir que o indivíduo se sinta apto a fazer trocas entre cenários de acordo com o que mais valoriza (Figura 13).

Figura 13 - Exemplo de aplicação de um par de cenários para escolha discreta



Fonte: Adaptado de Hjelmgren & Anell (2007). Population preferences and choice of primary care models: A discrete choice experiment in Sweden. *Health Policy*. 83, 314-322.

Assim, torna-se possível identificar o nível que influencia a escolha de cada atributo, i.e., estimar a taxa marginal de substituição dos atributos (Kjær 2005; Lancsar & Louviere 2008). O atributo custo (e.g., custos de transporte, salários, impostos) desempenha um papel extremamente importante na DCE na medida em que permite que o método proceda à elicitação das preferências relativamente à disposição para pagar dos indivíduos. Isto não significa que os respondentes sejam directamente questionados sobre a sua disponibilidade para pagar mas é-lhes pedido que efectuem trocas monetárias por melhorias nos atributos (Kjær 2005). De igual forma, o tempo de vida é também uma forma de medir o valor que se atribui a um estado de saúde.

De forma a analisar os resultados da DCE, procede-se à convergência do perfil de cada indivíduo respondente num único valor de utilidade, com atribuição de pesos a cada um dos atributos, constituindo uma combinação linear (Kjær 2005; Earl & Kemp 1999).

As escolhas na DCE analisam-se de acordo com a teoria da utilidade aleatória, i.e., para um indivíduo i condicionado à escolha j , a sua utilidade U pode ser desagregada numa componente explicável V_{ij} e numa componente não explicável, aleatória, ε_{ij} :

$$U_{ij} = V_{ij} + \varepsilon_{ij}, j = 1, \dots, J \quad [6]$$

Esta componente não explicável ε_{ij} pode ser devida a atributos não observáveis, variações não observáveis nas preferências ou resultado de erros de medição. A componente explicável V_{ij} é uma função dos atributos do bem ou serviço (vector X'_{ij} dos atributos) e das características dos indivíduos (vector Z'_i das características dos indivíduos):

$$V_{ij} = X'_{ij}\beta + Z'_i\gamma \quad [7]$$

Nesta função, β e γ representam os coeficientes a ser estimados. Assume-se que um indivíduo escolhe uma determinada opção 1, se e só se, a utilidade que lhe é atribuída é maior do que qualquer outra utilidade de uma opção num conjunto de J alternativas. Considerando que Y_i é uma variável aleatória que denota o resultado da escolha e assumindo uma distribuição de probabilidade conjunta para ε_i , a probabilidade P da utilidade ser maximizada ao escolher a opção 1 é dada por:

$$\begin{aligned} P(Y_i = 1) &= P(U_{i1} > U_{ij}) \\ &= P(V_{i1} + \varepsilon_{i1} > V_{ij} + \varepsilon_{ij}) \\ &= P(V_{i1} - V_{ij} > \varepsilon_{ij} - \varepsilon_{i1}), \forall j \neq 1 \end{aligned} \quad [8]$$

Não é prudente, no entanto, aplicar a DCE sem compreender na totalidade a técnica pelo que se torna fundamental estudar aprofundadamente a teoria, o desenho e

os resultados que se podem obter de forma a construir um modelo válido nos estudos de investigação (Lancsar & Louviere 2008).

Cada desenho de uma DCE tem em consideração o contexto específico da investigação em curso, dividindo-se por várias fases distintas com o intuito de simplificar a utilização do método. Vários autores defendem desenhos semelhantes ainda que possuam ligeiras discrepâncias entre si. No entanto, há algum consenso em dividir o desenho da DCE em cinco grandes fases principais: i) identificação dos atributos; ii) identificação dos níveis dos atributos; iii) desenho experimental; iv) recolha de dados; e v) análise de dados (Lancsar & Louviere 2008; Kjær 2005; Antunes et al. 2017). Para o sucesso da implementação da técnica é fundamental que os investigadores tenham plena noção do que está a ser estudado e que os respondentes tenham conhecimento do bem ou serviço em análise e sejam devidamente informados de forma a possibilitar escolhas assertivas (Lancsar & Louviere 2006).

A primeira fase do desenho da DCE respeita à identificação dos atributos. Estes devem ser abrangentes e mensuráveis, podendo ser do tipo qualitativo ou quantitativo. Para a identificação dos atributos mais adequados na caracterização de um bem ou serviço, há várias fontes de informação possíveis designadamente a literatura existente, grupos focais, entrevistas a pessoas chave (e.g., decisores) e peritos (Lancsar & Louviere 2008; Kjær 2005). O objectivo principal na definição e identificação dos atributos é que estes sejam relevantes para os decisores e que, simultaneamente, se revistam de significado para a maior parte dos respondentes. Embora não exista um limite máximo para o número de atributos, Kjær (2005) sugere não se exceder os oito atributos. Na identificação dos atributos há vários aspectos a ter em consideração nomeadamente se estes são irrelevantes ou não, i.e., até que ponto a sua exclusão/inclusão interfere com os resultados obtidos e se os atributos são, entre si, mutuamente dependentes ou possuem relação de causalidade. Estas características são particularmente importantes porque, se existirem, podem influenciar os indivíduos induzindo alguma diminuição da utilidade ou mesmo a transferência de utilidade entre atributos (Kjær 2005).

Finda a primeira fase, passa-se para a identificação dos níveis dos atributos. Assim, para cada um dos atributos identificados na etapa anterior são atribuídos níveis razoáveis

e relevantes. O intervalo de variação destes níveis deve ser tal que possibilite aos respondentes a realização de trocas entre as várias combinações de atributos. O número de níveis de um atributo influencia a sua importância, i.e., quantos mais níveis forem determinados para um atributo maior será a importância dada pelo respondente (em circunstâncias de intervalos de variação iguais). No entanto, quanto maior for o número e níveis dos atributos, mais complexo será o desenho da DCE. A inclusão de um atributo relacionado com o custo pode, segundo Kjær (2005) e Lancsar e Louviere (2008), ser ainda útil em alguns estudos nomeadamente quando se pretende efectuar análises económicas de custo-benefício.

Quando se pretende obter os valores de utilidade com recurso a instrumentos de medição já existentes e validados (e.g. EQ-5D e SF-6D), estas duas fases encontram-se naturalmente ultrapassadas dado que os atributos e respectivos níveis correspondem ao sistema descritivo destes instrumentos.

Na fase de desenho experimental procede-se à construção de várias combinações de atributos e conjuntos de escolhas. O objectivo principal é criar a DCE de tal forma que o número de alternativas seja o menor possível sem colocar em causa a capacidade do respondente inferir utilidade sobre todas as alternativas apresentadas (Kjær 2005). Este desenho vai permitir a estimação de uma matriz, resultante de todas as possíveis combinações dos níveis de atributos das distintas alternativas (conjunto de escolhas). O número de alternativas aumenta exponencialmente à medida que aumenta o número de atributos e o número de níveis associados. Assim, partindo da hipótese que todos os atributos têm o mesmo número de níveis, o número máximo de alternativas é dado por:

$$\text{número de alternativas} = \text{número de níveis}^{\text{número de atributos}}$$

[9]

Nas circunstâncias em que todas as alternativas possíveis são apresentadas aos respondentes está-se perante um desenho factorial completo. Este desenho, para além de garantir a independência dos efeitos relevantes dos atributos, permite estimá-los, bem como aos efeitos resultantes da interacção entre atributos. No entanto, por questões de

praticabilidade, o desenho factorial completo apenas pode ser aplicado a pequenas experiências, com um número muito limitado de atributos e de níveis. Em investigações com maior número de atributos e de níveis será, no entanto, necessário reduzir a dimensão do desenho da DCE e, nessas circunstâncias, o desenho factorial fraccionário constitui naturalmente a opção mais adequada. Nesta vertente mais reduzida utiliza-se um subconjunto ou fracção que garante as propriedades mais significativas do desenho factorial completo. Apesar de pressupor alguma perda de informação estatística, o desenho factorial fraccionário permite, tanto quanto possível, estimar eficientemente os efeitos dos atributos. Ainda que se considere o desenho factorial completo pouco praticável em estudos já com alguma dimensão, é sempre possível lidar com este tipo de desenho através de versões mais pequenas resultantes do modelo original. Com isto possibilita-se que sejam criados vários pontos de desenho sem aumentar o número do conjunto de escolhas para os respondentes (e.g., um desenho factorial completo com 256 combinações pode ser reduzido a 16 versões de 16 conjuntos de escolhas). À semelhança do que ocorre com o desenho factorial completo, a versão fraccionária também permite obter versões mais reduzidas. Importa assinalar que, apesar de no desenho factorial completo se utilizarem todas as alternativas possíveis, ainda assim, a eficiência do desenho de uma DCE necessita, para ser garantida, do correcto emparelhamento das alternativas na formação do conjunto de escolhas. Deste modo e de forma a assegurar a optimização do desenho é necessário que as seguintes propriedades se verifiquem: equilíbrio dos níveis (cada nível de um atributo surge com a mesma frequência nas alternativas incluídas no desenho); ortogonalidade (a diferença nos níveis de cada atributo varia de forma independente ao longo dos conjuntos de escolhas); justaposição mínima (os níveis não se repetem a si mesmo num conjunto de escolhas); e equilíbrio de utilidade (as utilidades das alternativas é aproximadamente igual em cada conjunto de escolhas). Os tipos de desenhos referidos podem ser obtidos através de classificações manuais ou preferivelmente com recurso a *software* específico (Kjær 2005; Lancsar & Louviere 2008).

Ultrapassada a fase de desenho experimental, procede-se à recolha de dados. Os estudos realizados sobre o desenho da DCE recomendam a inclusão de uma alternativa adicional no conjunto de escolhas, particularmente em circunstâncias em que é irrealista

considerar que o bem ou serviço é consumido com toda a certeza. Esta alternativa é considerada uma não escolha e impede que o indivíduo seja forçado a escolher algo que não é importante para si (Kjær 2005). Com a inclusão desta alternativa o respondente pode “optar por sair”, escolher “nenhuma”, “não participar” ou manter a situação existente (status quo). A limitação desta alternativa reside, em situações de modelos com alguma complexidade, no facto do indivíduo poder escolher a resposta mais fácil, i.e., aplicar uma heurística para evitar fazer uma escolha considerada difícil. Outra perspectiva relativa à inclusão de uma alternativa adicional é a possibilidade de esta reflectir uma situação de indiferença para o indivíduo (Kjær 2005; Lancsar & Louviere 2008). Para o sucesso da recolha de dados é fundamental garantir que o respondente compreenda o que se espera dele o que requer uma explicação exhaustiva por parte dos investigadores no que se refere ao contexto e objectivos do estudo, à descrição dos atributos e à contextualização de cada um dos cenários (conjuntos de escolhas). Nesta fase, os investigadores podem também considerar pertinente proceder à validação da técnica com recurso a testes de racionalidade tornando-se, para tal, necessário incluir conjuntos de escolhas adicionais para testar os diversos axiomas, tais como, integralidade, transitividade e continuidade, com defendem Lancsar e Louviere (2008). Por fim, é igualmente importante a inclusão de algumas perguntas relativas a dados sociodemográficos (e.g., idade, género, habilitações literárias, nível de rendimento, profissão). A recolha de dados pode ser efectuada com recurso a entrevistas, questionários, ou combinação destes métodos.

No que respeita à amostra esta deve reflectir a população para a qual os resultados vão ser generalizados e a quem os custos de oportunidade importam. A dimensão da amostra está muitas vezes sujeita ao orçamento da investigação e a outros eventuais constrangimentos, mas deverá sempre permitir a estimação de modelos com garantia de fiabilidade. A dimensão da amostra vai depender do número de perguntas a que cada indivíduo responde, do tamanho da população, da taxa de resposta que se espera obter no estudo e da potência estatística que se pretende obter do modelo (Lancsar & Louviere 2008; Kjær 2005).

A última fase respeita à análise de dados, que inclui a escolha do modelo de probabilidade e a inserção de dados, ambos com recurso a *software* apropriado, e também

a especificação do modelo estatístico. Segundo Kjær (2005), há vários modelos disponíveis para a estimação da DCE cuja escolha depende, entre outros factores, do desenho escolhido para o método. No caso da DCE ser dicotómica (resposta “sim/não” para uma determinada alternativa) ou apresentar um conjunto de escolhas constituído por duas alternativas são utilizados modelos binomiais de escolha discreta (*logit* ou *probit*). Quando o conjunto de escolhas é constituído por três ou mais alternativas é apropriada a utilização de modelos *logit* ou *probit* multinomiais, *mixed logit*, *nested logit* e valor extremo com heterocedasticidade. Este último conjunto de modelos, pelo aumento do número de alternativas no conjunto de escolha, coloca uma série de desafios que não existe na escolha binária (Ryan et al. 2008; Kjær 2005). Em geral, a estimação dos modelos de escolha discreta baseia-se no método da máxima verosimilhança, embora outros métodos possam também ser considerados (Kjær 2005). Os valores de utilidade obtidos através da DCE podem ser ancorados na escala de utilidade, que varia entre 0 e 1, incorporando a duração (tempo) como um atributo. Este método híbrido é designado por DCE_{TTO} (Bansback et al. 2014). Para além do *software* estatístico mais clássico como o SPSS ou o STATA, outros programas mais específicos, como o Lindep ou o Ngene, são também utilizados. A etapa de análise de dados pressupõe ainda a interpretação e a discussão dos resultados obtidos.

Os estudos comparativos entre a aplicação de diferentes técnicas para obtenção de valores de utilidade mostram que técnicas como a DCE são promissoras relativamente à utilização do TTO e do SG (Brazier et al. 2012; Krabbe et al. 2014). As vantagens identificadas na aplicação desta técnica residem na facilidade de compreensão por parte dos respondentes, permitindo a sua implementação a um maior número de pessoas, e por não ser afectada por preferências no tempo. Na realidade, a DCE tem vindo a obter cada vez mais destaque em estudos para eliciação de preferências em saúde (Antunes et al. 2017), existindo, actualmente, vários estudos publicados que utilizam a DCE como técnica de valoração de estados de saúde (Craig et al. 2013; Bansback et al. 2012; Xie et al. 2014; Shah et al. 2015; Robinson et al. 2015; Ratcliffe et al. 2011; Norman, Cronin, et al. 2013). Adicionalmente, a DCE tem vindo a ser aplicada em grupos particulares, como crianças e idosos (Ratcliffe et al. 2011; Ryan et al. 2006; Monteiro et al. 2014). No entanto, esta

técnica apresenta desafios, que requerem uma análise mais aprofundada, relacionados com a conversão de valores para a escala de utilidade (Rowen et al. 2015).

O futuro passará por efectuar estudos adicionais com recurso a esta técnica de elicitação de preferências e pela avaliação da implementação dos métodos híbridos que combinem a utilização de dados DCE (Bansback et al. 2014).

4. MEDIDAS DESCRITIVAS DE RESULTADOS EM SAÚDE

A avaliação económica envolvendo os QALY traduz o seu resultado sob a forma de custo por QALY. De forma a operacionalizar o cálculo dos QALY, há duas componentes a considerar: a primeira consiste na descrição do estado ou perfil de saúde de um indivíduo; a segunda representa a valoração dessas descrições. Para tal, torna-se necessário recorrer a instrumentos de medição em saúde, genéricos ou específicos, e baseados em preferências.

Os instrumentos de medição em saúde são considerados específicos quando são dirigidos, por exemplo, a uma patologia em particular (e.g. asma), a uma dimensão de saúde (e.g. dor muscular), ou a um grupo de indivíduos (e.g. idosos, crianças). Dado focarem-se em dimensões mais pertinentes para a condição em análise, os argumentos para a sua utilização incidem no facto de serem mais sensíveis na detecção de alterações no problema em estudo e mais relevantes para os interesses dos doentes. No entanto, podem não detectar o impacto de possíveis comorbilidades nos indivíduos, que podem ser mais relevantes que o problema em tratamento, e dos possíveis efeitos secundários das intervenções de saúde. Acresce ainda que os seus resultados não são comparáveis entre intervenções ou programas de saúde, associados a diferentes problemas. O principal argumento para a utilização, em avaliação económica, de instrumentos de medição em saúde genéricos e baseados em preferências reside no facto deste tipo de instrumentos permitir comparações entre programas de saúde que envolvem diferentes patologias e tratamentos. O NICE é uma das instituições que adopta este argumento, requerendo a utilização de um instrumento genérico baseado em preferências, nos estudos de referência de avaliação económica que elabora (Brazier et al. 2007). O mesmo sucede com o INFARMED, em Portugal, que, de acordo com as suas orientações metodológicas para estudos de avaliação de medicamentos aconselha a adopção de um instrumento genérico desde que validado para Portugal (INFARMED 1998).

Todos os instrumentos genéricos são compostos por um sistema descritivo e por algoritmos baseados em preferências, que garantem a obtenção de utilidade e que

agregam o valor desse sistema descritivo, generalizando-o para a sociedade. Embora todos estes instrumentos se baseiem em preferências, alguns investigadores têm detectado diferenças nos resultados, pelo facto de se basearem num sistema descritivo distinto (Ferreira, Ferreira & Pereira 2014).

São vários os instrumentos genéricos de medição em saúde baseados em preferências, destacando-se cinco com maior utilização: *Quality of Well-Being* (QWB), 15-D, *Health Utilities Index* (HUI), SF-6D e EQ-5D.

4.1. Quality of Well-Being

O QWB, desenvolvido na década de 1970, foi o primeiro instrumento de medição da qualidade de vida desenhado especificamente para a estimação de QALY (Kaplan et al. 1976). Este instrumento, utilizando a VAS como técnica de valoração, permite a ponderação das preferências dos indivíduos através da combinação de três escalas funcionais multinível (mobilidade, actividade física e actividade social) com uma medida de aversão a sintoma-problema (em caso de múltiplos sintomas ou problemas, aquele que o indivíduo identifica como o menos desejável, será o sintoma-problema utilizado). O objectivo é produzir, para um momento no tempo, um valor (e não utilidades) que se situe numa escala de preferências que varia entre 0 e 1, ou morte e saúde perfeita (Seiber et al. 2008; Drummond et al. 2005).

4.2. Health Utilities Index

O HUI corresponde a uma família de instrumentos de medição em saúde genéricos e baseados em preferências, cujo intuito é a obtenção de utilidades (Horsman et al. 2003). A primeira versão (HUI1) foi desenvolvida por Torrance e colaboradores nos finais da década de 1970 para utilização em avaliação económica de cuidados intensivos neonatais (Brazier et al. 2007). A esta versão seguiram-se duas outras – sistemas HUI2 e HUI3 – que, no seu conjunto, descrevem quase um milhão de estados de saúde. Na maior parte dos

casos, a abordagem inicial é feita com a medida HUI3, pelo facto de possuir um sistema descritivo mais detalhado, ter uma estrutura independente e possuir normas validadas, generalizadas à população. A medida HUI2 possui alguns atributos adicionais relativamente à HUI3, sendo mais apropriada para estudos específicos. A obtenção de utilidades é efectuada através da técnica do SG, sendo balizada por uma escala que varia entre 0 ou morte e 1 ou saúde perfeita (Horsman et al. 2003; Drummond et al. 2005).

4.3. 15-D

O 15-D é um instrumento genérico, constituído por 15 dimensões, que pode ser auto-aplicado por indivíduos adultos, na medição da qualidade de vida relacionada com a saúde. O valor obtido com a aplicação do 15-D tem por base um conjunto de pesos de utilidades ou de preferências e situa-se numa escala que varia entre 0 ou morte e 1 ou saúde perfeita. A obtenção desse valor é efectuada através da utilização combinada das técnicas de escala de pontuação e de estimação de magnitude (Sintonen 1995; Sintonen 1994).

4.4. EQ-5D

O EQ-5D é um instrumento genérico de medição em saúde, baseado em preferências, desenvolvido em 1988 pelo grupo EuroQol que integra um conjunto de investigadores oriundos de diversos países (EuroQol Group, 1990).

Este instrumento deriva de um sistema descritivo inicial com seis atributos (mobilidade, cuidados pessoais, actividades habituais, relações sociais, dor/mal-estar e ansiedade/depressão), revisto em 1990 para incluir apenas cinco atributos, excluindo-se as relações sociais (EuroQol Group 2005). No desenvolvimento deste sistema descritivo, os autores pretenderam cumprir os seguintes requisitos: i) conter dimensões de saúde significativas para doentes e para a população em geral; ii) ser simples (poucas dimensões com o menor número de níveis possível); iii) possuir uma descrição curta e clara dos

estados de saúde de forma a possibilitar a distinção entre eles; e iv) de preenchimento amigável em contextos distintos (Brooks, 2013). A versão original deste instrumento (EQ-5D-3L) determinava três níveis para cada atributo – nenhum problema, alguns problemas e problemas extremos – possibilitando a definição de 243 (3^5) possíveis estados de saúde. O EQ-5D-3L revelou-se um instrumento com utilidade em diversas condições de saúde e na obtenção de valores da população em geral, encontrando-se traduzido para mais de 100 línguas.

A sua utilização revelou, no entanto, que pode sofrer de efeito-tecto, quando utilizado na população em geral e em alguns grupos de doentes, traduzido na falta de sensibilidade em medir pequenas alterações nos estados de saúde mais ligeiros. Isto significa que o EQ-5D-3L detém uma característica que dificulta a distinção dos estados de saúde situados na parte superior da escala de utilidade (Bharmal & Thomas 2006; Brazier et al. 2004; Cunillera et al. 2010; EuroQol Group 2014; Ferreira, Ferreira & Pereira 2014; Herdman et al. 2011; Petrou & Hockley 2005). Esta particularidade faz com que o sistema de valores do EQ-5D-3L origine valores de utilidade mais elevados nos indivíduos em melhor situação de saúde, comparativamente com os valores obtidos pelo algoritmo do SF-6D para os mesmos indivíduos (Cunillera et al. 2010; Kontodimopoulos et al. 2009). Esta limitação conduziu a uma investigação exaustiva por parte do Grupo EuroQol e à apresentação, em 2009, de um instrumento de medição mais sensível, o EQ-5D-5L, que mantém as dimensões originais mas aumenta para cinco o número de níveis dentro de cada dimensão – nenhum problema, problemas ligeiros, problemas moderados, problemas graves e problemas extremos – permitindo a descrição de 3.125 (5^5) estados de saúde (Herdman et al. 2011).

A técnica de valoração utilizada para obter o sistema de valores no EQ-5D-3L tem sido, maioritariamente, a técnica TTO. No entanto, dada a complexidade do TTO, o grupo EuroQol decidiu explorar o potencial de outras técnicas de valoração, designadamente da DCE. O objectivo foi determinar se esta técnica podia providenciar informação adicional que complementasse os valores obtidos através do TTO (Oppe et al. 2014). Neste âmbito, o grupo EuroQol introduziu a DCE como técnica de valoração no EQ-5D num estudo-piloto realizado na Holanda, em 2008. Não obstante os valores obtidos através da técnica DCE

serem ligeiramente superiores, os resultados mostraram que, de uma forma geral, estes replicavam os valores tidos com recurso ao TTO (Stolk et al. 2010). No entanto, a utilização desta técnica de valoração não é isenta de crítica. Isso deve-se sobretudo à dificuldade em ancorar os valores gerados pela DCE na escala de utilidade cuja métrica se situa entre 0 (morte) e 1 (saúde perfeita). Apesar destas limitações, dados os resultados preliminares obtidos o grupo EuroQol decidiu aprofundar esta linha de investigação ao incluir a tarefa DCE, a par com a tarefa TTO, nos estudos de valoração do EQ-5D-5L (Stolk et al. 2010).

Adicionalmente, o grupo EuroQol tem desenvolvido uma investigação aprofundada sobre várias abordagens das variantes *lead-time* e *lag-time* do TTO (Oppe et al. 2014).

A utilização de modelos econométricos para os dados obtidos através das técnicas de valoração permite a obtenção dos valores de utilidade associados a cada estado de saúde. O valor vai situar-se numa escala que varia entre 0 ou morte a 1 ou saúde perfeita, permitindo, no entanto, valores negativos correspondentes a estados de saúde considerados piores do que morte (P. Ferreira et al. 2013; Ferreira 2010; Drummond et al. 2005).

A versão portuguesa do questionário EQ-5D-3L foi finalizada em 1998 e publicada, em 2013, por P. Ferreira et al. (2013). Para além da elaboração da versão portuguesa, foi ainda estimado, numa amostra representativa da população portuguesa, o sistema de valores para o EQ-5D-3L com recurso ao TTO e definidas as normas populacionais para Portugal (Ferreira, Ferreira, Pereira, et al. 2014b; Ferreira, Ferreira, Pereira, et al. 2014a). Em 2015, o CEISUC iniciou um projecto de investigação, em parceria com o grupo EuroQol, com o intuito de obter o sistema de valores para a população portuguesa recorrendo ao EQ-5D-5L. Os resultados deste estudo consubstanciam o objectivo da presente tese.⁸

A Figura 14 apresenta o questionário com as cinco dimensões e respectivos cinco níveis do EQ-5D-5L na versão portuguesa.

⁸ Embora os resultados não tenham ainda sido publicados em revistas de cariz científico, estão, neste momento, a ser preparados para submissão a curto prazo, tendo sido já apresentados resultados preliminares em diversas conferências nacionais e internacionais de economia da saúde e qualidade de vida.

Figura 14 - Versão portuguesa do questionário EQ-5D-5L

MOBILIDADE

- Não tenho problemas em andar
- Tenho problemas ligeiros em andar
- Tenho problemas moderados em andar
- Tenho problemas graves em andar
- Sou incapaz de andar

CUIDADOS PESSOAIS

- Não tenho problemas em me lavar ou vestir
- Tenho problemas ligeiros em me lavar ou vestir
- Tenho problemas moderados em me lavar ou vestir
- Tenho problemas graves em me lavar ou vestir
- Sou incapaz de me lavar ou vestir

ATIVIDADES HABITUAIS (ex. trabalho, estudos, atividades domésticas, atividades em família ou de lazer)

- Não tenho problemas em desempenhar as minhas atividades habituais
- Tenho problemas ligeiros em desempenhar as minhas atividades habituais
- Tenho problemas moderados em desempenhar as minhas atividades habituais
- Tenho problemas graves 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 ligeiros
- Tenho dores ou mal-estar moderados
- Tenho dores ou mal-estar graves
- Tenho dores ou mal-estar extremos

ANSIEDADE/DEPRESSÃO

- Não estou ansioso/a ou deprimido/a
- Estou ligeiramente ansioso/a ou deprimido/a
- Estou moderadamente ansioso/a ou deprimido/a
- Estou gravemente ansioso/a ou deprimido/a
- Estou extremamente ansioso/a ou deprimido/a

Portugal (Portuguese) © 1997 EuroQol Group. EQ-5D™ is a trade mark of the EuroQol Group

4.5. SF-6D

O SF-6D é um instrumento de medição, desenvolvido a partir do conhecido questionário SF-36 *Health Survey* ou SF-36 (Brazier et al. 2002; Brazier et al. 1998). Dado o SF-36 ter um uso limitado em avaliação económica, pelo facto dos seus resultados não se

converterem em utilidades e, por conseguinte, em QALY, surgiu a necessidade de desenvolver o algoritmo SF-6D permitindo, assim, traduzir os resultados dos estados de saúde do SF-36 em utilidades. Este algoritmo deriva de 11 itens do SF-36 (oito provenientes da versão reduzida do SF-36, o SF-12, e três provenientes do próprio SF-36), combinados em seis dimensões de saúde (função física, limitação no desempenho, função social, dor, saúde mental e vitalidade), cada uma com quatro a seis níveis, permitindo a criação de 18.000 estados de saúde distintos (Brazier et al. 2002). A utilização de um modelo econométrico permite a atribuição de pesos a cada uma das dimensões do SF-6D e, conseqüentemente, a atribuição de valores (numa escala de 0 ou morte a 1 ou saúde perfeita) para os estados de saúde definidos. Em qualquer análise, a utilização do SF-6D não pode ser efectuada de forma independente, devendo ser precedida pela aplicação do questionário SF-36, de forma a poderem ser obtidas utilidades (L. Ferreira et al. 2013). Para obtenção do sistema de valores recorre-se à técnica de medição de utilidades SG (Brazier et al. 2007; Drummond et al. 2005).

Os trabalhos de investigação realizados nos últimos anos com recurso ao SF-6D têm proposto algumas melhorias a este instrumento de medição, nomeadamente no que concerne ao seu sistema descritivo e ao seu sistema de valores. Embora nem todos os estudos o demonstrem, é geralmente reconhecido que o SF-6D sofre de um efeito-chão, i.e., uma característica que dificulta a descrição dos estados de saúde situados na parte inferior da escala de utilidade (Brazier et al. 2004; Ferreira et al. 2011). Esta particularidade faz com que o sistema de valores do SF-6D origine um valor mínimo mais elevado do que o menor valor gerado pelo algoritmo do EQ-5D, resultando em utilidades mais elevadas nos indivíduos em pior situação de saúde. Uma forma de contornar esta limitação do SF-6D pode consubstanciar-se em alterações ao seu sistema descritivo, nomeadamente no aumento do número de dimensões e/ou aumento dos níveis em cada dimensão, bem como na procura de alternativas à forma de obtenção do sistema de valores.

Na verdade, já Ferreira et al. (2011) haviam proposto uma alteração ao sistema descritivo do SF-6D, com o intuito de mitigar a limitação do efeito-chão, adicionando um nível às dimensões “função física” e “limitação no desempenho”. Os resultados do estudo destes autores sugerem que a alteração do sistema descritivo pode ser um caminho a

seguir para resolver a limitação do SF-6D mas que há necessidade de investigação adicional. Outros estudos foram entretanto publicados sobre o mesmo problema, como o de Abellán Perpiñán e colegas (Abellán Perpiñán et al. 2012).

A versão portuguesa do SF-6D foi publicada em 2006 (Ferreira & Ferreira 2006), tendo sido publicado, em 2010, o respectivo sistema de valores (Ferreira, Ferreira, et al. 2010).

As normas do SF-6D por variáveis sociodemográficas foram também obtidas para uma amostra representativa da população portuguesa e publicadas em 2015 (Ferreira et al. 2015).

Adicionalmente, foram ainda efectuados alguns estudos de âmbito nacional no sentido de validar o SF-6D em amostras de indivíduos com doenças de elevada prevalência no país tais como artrite reumatóide, asma, catarata e doença pulmonar obstrutiva crónica (Ferreira, Ferreira & Pereira 2014; Ferreira, Brito, et al. 2010; Ferreira & Ferreira 2008; Ferreira et al. 2008).

4.6. Fontes de Informação

Finda a apresentação dos instrumentos de medição em saúde genéricos mais utilizados, importa ainda analisar a questão sobre quem deve valorar o respectivo sistema descritivo ou os estados de saúde, isto é, a quem devem ser elicitadas as preferências?

A valoração em saúde pode ser obtida através de diversas fontes: i) grupos específicos como doentes, cuidadores ou profissionais de saúde; ii) ou, numa perspectiva mais abrangente, elementos da comunidade. O Quadro 5 resume as diferentes abordagens possíveis no que respeita a quem deve ser questionado sobre as preferências relativas aos estados de saúde.

Quadro 5 - Diferentes abordagens para a obtenção de preferências

Objectivo	Decisão clínica pessoal	Auscultação societária: avaliação de programas ou actividades de saúde em curso	Alocação societária de recursos: definição de prioridades em programas de saúde
Quem questionar	O indivíduo, informado por outros doentes ou pessoas com a incapacidade	Os indivíduos afectados por essa actividade (e.g., doentes ou pessoas com incapacidade, indivíduos que evitaram a doença)	Amostra representativa da população

Fonte: Adaptado de Weinstein et al. (2009). QALYs: The basics. Value in Health. 12 (Supl. 1), 5-9, Tabela 1

Porque é que esta questão é importante? Porque o tipo de indivíduos que valora os estados de saúde pode influenciar os resultados obtidos. Brazier et al. (2007) referem que estas discrepâncias ocorrem essencialmente devido a descrições pobres dos estados de saúde, providenciadas no sistema descritivo, e a alterações dos padrões pessoais devidas a mudanças na saúde dos indivíduos (e.g. quando um idoso valora a sua saúde em função do que considera ser o melhor estado de saúde expectável para um indivíduo desse grupo etário e não o melhor estado de saúde por si).

Numa perspectiva societária, a decisão sobre a alocação de recursos escassos pode assentar em valores obtidos através de indivíduos que, no momento da valoração, experienciam ou experienciaram o estado de saúde de interesse, ou através de indivíduos que tomam decisões por estes, valorando um estado de saúde que podem ou não ter vivenciado (Brazier et al. 2007; Weinstein et al. 2009). Esta questão é particularmente relevante e promotora de alguma discussão sobre se a valoração deve concentrar-se na população em geral (i.e. baseada em preferências) ou em doentes (i.e. baseada em experiência). A fronteira entre estas duas abordagens é, no entanto, algo ténue na medida em que aos indivíduos que não experienciaram estados de saúde em particular poderá ser disponibilizado conhecimento prévio, através de doentes ou pessoas com incapacidade, sobre o que significa vivenciar esse estado. Desta forma, a abordagem *ex ante* (i.e. antes de qualquer experiência prévia) pode ser enriquecida com a informação experienciada por outros, não sendo necessária a experiência pelo estado de saúde (*ex post*) para o valorar (Weinstein et al. 2009). Por outro lado, indivíduos que tenham experienciado uma doença

ou incapacidade podem criar estratégias de adaptação, a nível físico e emocional, condicionando a sua valoração dos estados de saúde (Brazier et al. 2007).

No entanto, no que concerne ao contexto das decisões individuais predominam claramente as preferências do indivíduo de interesse, desde que a pessoa esteja devidamente informada por outros com experiência no estado de saúde que o indivíduo pode vir a experimentar (Weinstein et al. 2009).

No que concerne ao modelo convencional do QALY, por exemplo, os seus utilizadores defendem que os valores devem basear-se nas preferências obtidas *ex ante* através de uma amostra representativa da população em geral (Lipscomb et al. 2009). Um forte argumento é o véu de ignorância que a comunidade possui quando valora, alheia ao seu interesse pessoal e a qualquer estado de saúde em particular (Brazier et al. 2007).

5. MATERIAIS E MÉTODOS

Os trabalhos de investigação realizados nos últimos anos com recurso ao EQ-5D têm proposto algumas melhorias a este instrumento de medição, nomeadamente no que concerne ao seu sistema descritivo e ao seu sistema de valores. Por outro lado, o reconhecimento da complexidade da técnica TTO para obtenção do sistema de valores conduziu à introdução de outras técnicas de valoração, como a DCE, para obtenção de informação complementar à já adquirida com o TTO.

No que concerne ao sistema de valores, importa que os dados obtidos tenham significado e relevância para o decisor, i.e., que reflectam o contexto para o qual a decisão sobre a alocação de recursos está a ser tomada (EuroQol Group 2007). Os resultados de vários estudos vão ao encontro de que variáveis de contexto, como a nacionalidade, parecem influenciar os valores obtidos através da técnica TTO (Andrade et al. 2013; Busschbach et al. 2003; Huang et al. 2007; Johnson et al. 2005; Sakthong et al. 2008). Isto sugere que há diferenças entre as populações no que concerne à valoração da saúde. Esta diferença pode ser resultante das características demográficas e de aspectos culturais e tal conclusão implica que, desde que disponível, se utilize um sistema nacional de valores obtido a partir das preferências da população de cada país (Tsuchiya et al. 2002; Jelsma et al. 2003; Johnson et al. 2005; Yusof et al. 2012).

Após este sucinto enquadramento, apresenta-se, neste capítulo, os objectivos deste estudo, as questões subjacentes ao trabalho de investigação e os procedimentos metodológicos que permitiram a sua implementação e concretização.

5.1. Objectivo

O objectivo principal deste estudo foi estimar o sistema de valores dos estados de saúde descritos pelo EQ-5D-5L para Portugal. Para tal, utilizou-se a versão portuguesa do instrumento genérico de medição em saúde EQ-5D-5L já utilizada previamente no estudo

realizado por Ferreira et. al (2016), elicitaram-se as preferências de uma amostra representativa da população portuguesa e estimaram-se diversos modelos econométricos com o intuito de obter o sistema de valores de todos os estados de saúde descritos por este instrumento de medição, para Portugal.

5.2. Questão de Investigação

Este trabalho de investigação pretendeu averiguar qual o modelo econométrico que melhor se adequa à estimação do sistema de valores do EQ-5D-5L para a população portuguesa, i.e., qual o modelo que melhor representa as preferências nacionais de forma a suportar os processos de tomada de decisão em saúde, avaliação económica, ganhos em saúde e políticas em saúde.

5.3. Procedimentos

Os pontos seguintes apresentam, de forma cronológica, as etapas preparatórias que precederam a implementação do estudo.

5.3.1. Apresentação da candidatura

O projecto iniciou-se, em Novembro de 2014, com a apresentação por parte dos investigadores nacionais de uma candidatura ao grupo EuroQol, que detém os direitos de propriedade intelectual do instrumento de medição em saúde EQ-5D, para a realização de um estudo de âmbito nacional com vista a determinar o sistema de valores para os 3.125 estados de saúde descritos pelo EQ-5D-5L, numa amostra representativa da população portuguesa.

O EuroQol é uma organização sem fins lucrativos, constituída por uma rede internacional de mais de 75 investigadores, que tem por objectivo melhorar o processo de decisão na área da saúde ao promover a utilização de instrumentos de medição em saúde com aplicabilidade mundial.

A candidatura consistiu numa apresentação detalhada do plano de trabalhos proposto: contextualização do problema, pertinência e objectivos do estudo, questões de investigação, técnicas de valoração a utilizar, caracterização e dimensão da amostra, métodos para recolha de dados, análise e modelação de dados e principais resultados esperados. Adicionalmente, foi proposto, pelo grupo coordenador nacional, um cronograma para a concretização das etapas previstas na prossecução do projecto e estimado um orçamento para o suporte da recolha de dados.

Em Junho de 2015, o grupo EuroQol autorizou a realização do projecto nos moldes propostos, tendo firmado com o grupo coordenador nacional o compromisso de adesão a um código de conduta associado aos estudos de valoração do EQ-5D-5L.

5.3.2. Código de conduta

Os resultados obtidos por estudos anteriores de valoração do EQ-5D-5L, realizados em diversos países, careciam de comparabilidade em parte devido à heterogeneidade dos protocolos de investigação seguidos (Knies et al. 2009). De forma a ultrapassar esta limitação, o grupo EuroQol definiu um código de conduta a ser seguido nos estudos de valoração do EQ-5D-5L pelas equipas responsáveis pela investigação em cada país.

Este código especifica a articulação entre o grupo coordenador nacional, que propõe a realização do estudo, e as equipas de suporte do grupo EuroQol e determina a adesão da equipa nacional ao que se designa por protocolo EQ-VT. Com esta normalização, o que se pretende é aumentar a comparabilidade dos estudos internacionais de valoração e aumentar a probabilidade de que as diferenças observadas nos valores obtidos para os estados de saúde do EQ-5D-5L se devam essencialmente às preferências locais, em função das características das populações, e não a diferenças metodológicas entre os estudos. Com este protocolo de valoração, que resulta do trabalho desenvolvido por vários estudos empíricos, estabelecem-se os critérios comuns para a recolha de dados em estudos desta natureza (Oppe et al. 2016; Versteegh et al. 2016).

Para além do protocolo de valoração, o código de conduta determina ainda um conjunto de requisitos respeitante ao número mínimo de indivíduos a incluir na amostra de cada estudo de valoração (mínimo de 1.000 participantes), à forma de recolha de dados

(entrevistas assistidas por computador ou CAPI) e ao número e treino dos entrevistadores.⁹ No que concerne aos entrevistadores e considerando a amostra prevista, o grupo EuroQol define um número ideal de oito até 14 entrevistadores, com vista à realização de 70 a 130 entrevistas cada. Ao garantir um número mínimo de entrevistadores, pretende-se atenuar o efeito individual do entrevistador na condução das entrevistas e, conseqüentemente, nos dados obtidos. O número de entrevistadores poderá, no entanto, ser ultrapassado desde que autorizado pela equipa de suporte do EQ-VT (grupo EuroQol) mas as equipas nacionais assumem que essa situação pressupõe uma dificuldade acrescida no acompanhamento do desempenho dos entrevistadores. O código de conduta especifica também a designação de um investigador na equipa nacional, com o perfil de superutilizador, como primeiro ponto de contacto com os entrevistadores no apoio à realização das entrevistas e ligação à equipa de suporte do EQ-VT, em caso de necessidade de apoio adicional na resolução de problemas. No código de conduta definem-se ainda os critérios para a realização e avaliação do teste piloto por cada entrevistador, antes de se proceder à recolha real de dados, é explicitada a necessidade de adesão dos entrevistadores ao guião previsto para a realização das entrevistas, e estabelecem-se os critérios mínimos de qualidade dos dados de cada entrevista. Na eventualidade dos entrevistadores não aderirem ao protocolo estabelecido, o código prevê os passos a seguir, nomeadamente o seu retreino e, eventualmente, o abandono do estudo. De forma a facilitar o controlo de qualidade, o código de conduta define ainda paragens na recolha de dados em etapas pré-estabelecidas (200, 400 e 600 entrevistas completas) nas quais a equipa de suporte do EQ-VT prevê a reunião com a equipa nacional com o intuito de discutir o desempenho e a adesão dos entrevistadores bem como a qualidade da informação. Prévio ao início do estudo, o código de conduta estabelece a realização de uma reunião preparatória entre a equipa de suporte do EQ-VT e o grupo coordenador nacional relativamente ao planeamento da implementação do estudo, à estratégia de recrutamento de investigadores e ao controlo da qualidade e adesão ao protocolo. Por fim, define o compromisso por parte da equipa nacional de publicação dos resultados obtidos em revistas científicas com reconhecimento internacional.

⁹ CAPI – *Computer assisted personal interviews*

5.3.3. Planeamento

Antes de se proceder à implementação do estudo a nível nacional, foi necessário realizar um conjunto de tarefas preliminares já previstas no código de conduta explicitado anteriormente.

Em Junho de 2015, o grupo coordenador nacional procedeu à primeira reunião com a equipa de suporte do EQ-VT de forma a definir uma calendarização relativamente ao planeamento da implementação do estudo, à estratégia de recrutamento de investigadores, ao controlo da qualidade dos dados e à adesão ao protocolo estabelecido pelo grupo EuroQol. O cronograma previu um prazo de nove semanas desde o momento da reunião preparatória até ao início da recolha de dados.

Após a definição dos prazos, o grupo EuroQol disponibilizou uma primeira versão, em inglês, do *software* por si desenvolvido para a recolha de dados: EQ-VT versão 2.0. Neste âmbito, a equipa nacional efectuou a tradução, validação e adaptação do conteúdo dos três perfis de acesso à plataforma (entrevistador, coordenador dos dados e investigador principal) e de todos os ecrãs constituintes desses perfis, com particular ênfase na entrevista guiada por computador e em possíveis informações geradas pelo sistema no decorrer de uma entrevista. Relativamente à tradução destes componentes, esta foi também validada por uma especialista em linguística.

No que concerne ao sistema descritivo do EQ-5D-5L, a equipa nacional recorreu à versão previamente utilizada no estudo implementado por Ferreira et. al (2016).

A equipa nacional foi ainda responsável pela criação de um conjunto de questões de âmbito local para inclusão na plataforma EQ-VT como questionário adicional, após a aplicação do questionário principal.

Para além disso, a equipa nacional colaborou na tradução e adaptação do guião das entrevistas a disponibilizar a cada entrevistador participante no estudo.

De forma a preparar o grupo coordenador nacional para a implementação do estudo, a equipa de suporte do EQ-VT providenciou uma formação com a duração de um dia, destinada à equipa nacional de investigadores, que decorreu no início de Setembro de 2015, nas instalações do EuroQol, em Roterdão, Holanda. Este treino incidiu sobre vários aspectos. Em primeiro lugar, visou a identificação de outros estudos de valoração em saúde

com recurso ao EQ-5D-5L que tivessem decorrido ou estivessem ainda em curso a nível internacional (Figura 15).

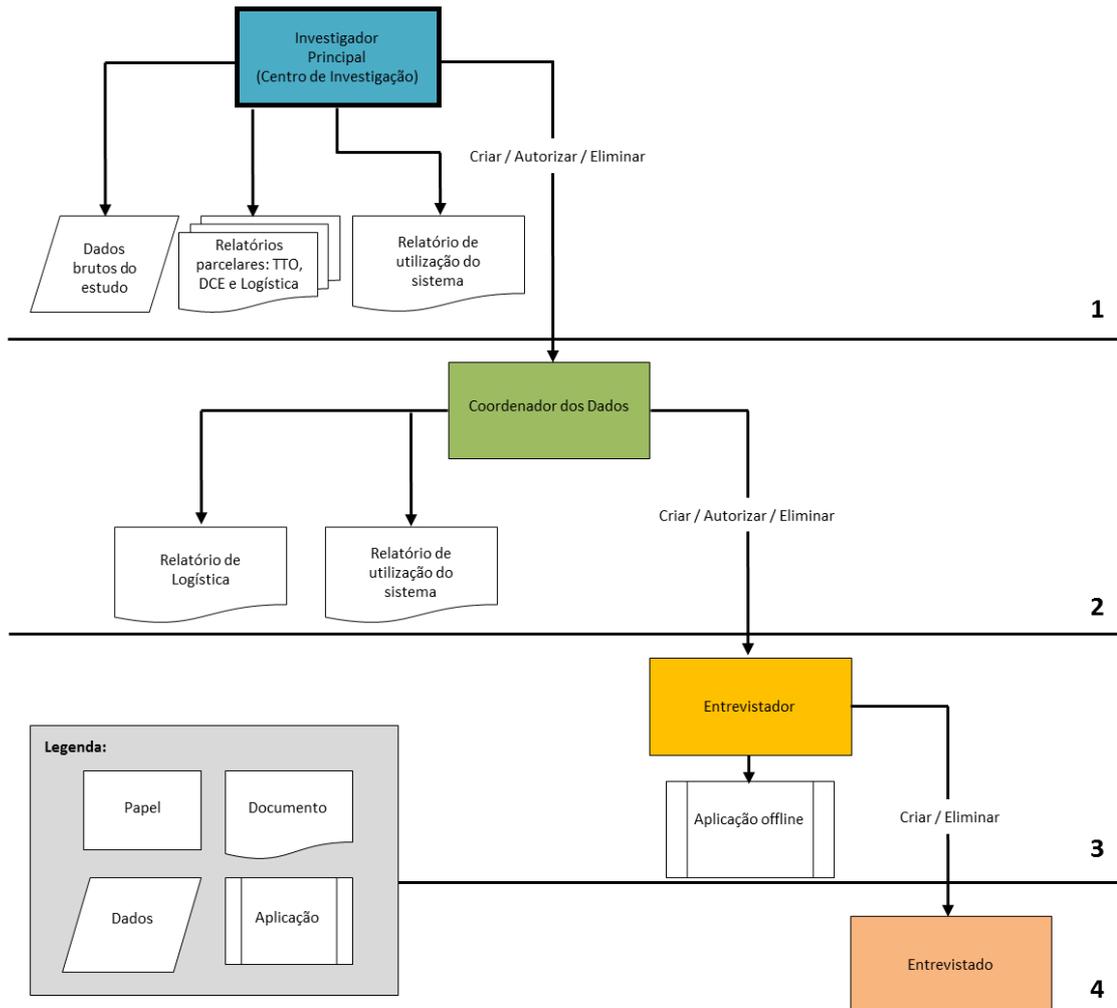
Figura 15 - Países com estudos de valoração do EQ-5D-5L concluídos ou em curso (Fevereiro de 2017)



Em segundo lugar, esta sessão destinou-se à apresentação do *software* EQ-VT e dos três níveis de acesso existentes. Em função do acesso escolhido, é possível aceder a um conjunto distinto de funções, tais como: i) realizar uma entrevista, ii) criar credenciais de acesso para novos entrevistadores e compilar informação de entrevistadores inscritos e iii) descarregar os ficheiros parcelares dos dados para análise. A Figura 16 evidencia a hierarquia de perfis existente e as respectivas funções acedidas. O treino incidiu de seguida sobre a ferramenta de controlo de qualidade QC Tool (*Quality Control Tool*) que foi criada pelo grupo EuroQol com o intuito de avaliar e acompanhar o desempenho dos entrevistadores, particularmente a sua adesão ao protocolo EQ-VT. Em subcapítulo posterior proceder-se-á a uma explicação mais detalhada sobre esta ferramenta de controlo de qualidade. Uma das particularidades da ferramenta QC Tool é a produção de

relatórios globais ou parcelares para análise dos dados recolhidos. Neste âmbito, a formação incidu também na interpretação dos resultados do relatório. Por fim, o treino incidu na identificação de eventuais problemas da plataforma EQ-VT durante a sua utilização e respectivas soluções.

Figura 16 - EQ-VT: hierarquia de perfis e respectivas funções do utilizador



5.3.4. Amostra e recolha de dados

A população alvo definida pelos investigadores, para este estudo, foi a população portuguesa com idade igual ou superior a 18 anos. De acordo com os resultados dos últimos Censos, existiam, em 2011, 8.657.240 indivíduos a residir em Portugal que obedeciam a este critério (Instituto Nacional de Estatística 2012).

Não obstante o EuroQol não definir qualquer obrigatoriedade sobre a amostra ser representativa da população nacional em geral, o grupo coordenador nacional decidiu seleccionar uma amostra aleatória e representativa da população portuguesa no que concerne ao género e idade.

Tal como referido anteriormente, o grupo EuroQol recomenda que, nos estudos de valoração do EQ-5D-5L, a amostra seja constituída por um número igual ou superior a 1.000 indivíduos. Este número advém da existência de requisitos específicos relativamente à estimação do valor médio do cTTO. Os cálculos efectuados por Ramos-Goñi et al. (2014) estimam que para obter um desvio-padrão de 0,01 em relação à média observada do valor do cTTO são necessárias 9.735 respostas. Como cada participante valora, durante a entrevista, 10 estados de saúde com recurso à técnica cTTO, tal conduz à necessidade de existirem pelo menos 1.000 indivíduos a integrar a amostra, de forma a obter 10.000 valorações.

Para a obtenção desta amostra, os investigadores delinearão duas estratégias: em primeiro lugar, estabeleceram o recurso a uma empresa de sondagens, especialista em recolha de dados em entrevistas porta-a-porta e com experiência prévia na valoração do EQ-5D-3L, com vista ao recrutamento de entrevistadores para a realização de 700 entrevistas na zona de Lisboa e Évora; em segundo lugar, concretizaram um processo directo de recrutamento de entrevistadores em Coimbra, também com o objectivo inicial de realizar 700 entrevistas.

De forma a obter uma amostra representativa da população portuguesa, foi previamente definida a distribuição da amostra pelo grupo de investigadores considerando estratos por género (2 classes) e grupo etário (4 classes), proporcionais à distribuição da população portuguesa segundo o último Censos. Os cálculos iniciais efectuados tiveram por base a dimensão mínima da amostra de 1.000 indivíduos, sendo obviamente proporcionalmente ajustados para amostras superiores. A tabela 1 apresenta a projecção da amostra tendo por base a estratificação por género e grupo etário.

Tabela 1 - Projecção da amostra representativa da população portuguesa, estratificada por género e grupo etário

Grupo Etário	Género	Masculino	Feminino	Total
18-29 anos		79	78	157
30-49 anos		169	182	352
50-69 anos		147	167	314
>=70 anos		71	107	178
Total		467	533	1.000

A amostra aleatória estratificada por género e grupo etário foi recolhida com base no método de amostragem do itinerário aleatório, seguindo as seguintes etapas: i) selecção aleatória de uma freguesia; ii) selecção aleatória do ponto de partida (com base em mapas); iii) a partir de cada ponto de partida, determinação de um caminho quase aleatório, com base num esquema predefinido (esquerda ou direita relativamente ao passeio; regras para cruzamentos e becos sem saída); iv) selecção dos edifícios (1.º, 3.º, 5.º edifícios,...); e v) selecção do indivíduo a inquirir na habitação (critério da próxima pessoa do agregado a fazer anos).

Prévia à recolha de dados, programou-se uma formação com os potenciais entrevistadores, sob a responsabilidade do grupo de investigadores. A formação, com carácter individual ou de grupo, foi organizada com uma duração média de quatro horas e com a seguinte estrutura: apresentação dos objectivos do estudo e da importância deste tipo de trabalhos de investigação no contexto da tomada de decisão; apresentação do instrumento genérico de medição em saúde EQ-5D-5L e explicação das suas cinco dimensões e níveis; apresentação da plataforma EQ-VT e explicitação dos acessos; simulação completa de uma entrevista; e forma de resolução de eventuais problemas com a utilização da plataforma EQ-VT durante a realização das entrevistas.

Finda a formação inicial, estabeleceu-se a realização, por cada potencial entrevistador, de um teste piloto composto por cinco entrevistas. Se no teste piloto associado a cada entrevistador, duas ou mais entrevistas (40% ou superior) falhassem os critérios principais de qualidade previstos no código de conduta apresentados previamente a cada entrevistador e explicitados em subcapítulo posterior, o entrevistador teria nova

formação e realizaria um novo teste piloto. Se ainda assim falhasse o controlo de qualidade, o entrevistador seria forçado a abandonar o estudo. Caso os critérios principais de qualidade fossem garantidos, os entrevistadores obtinham a autorização dos investigadores para o início da recolha de dados. Importa ainda salientar que para a realização do trabalho de campo, os entrevistadores dispunham de uma carta de apresentação do estudo para disponibilização aos indivíduos entrevistados, informação actualizada sobre o número necessário de participantes relativamente ao género e grupo etário e orientações sobre o local a realizar as entrevistas. No que concerne à empresa de sondagens, acresce ainda que após a selecção dos entrevistadores de acordo com os requisitos do estudo, estes assistiam a uma formação adicional sobre regras de condução de entrevistas pessoais, ficando adicionalmente munidos de um manual do entrevistador e contando com o apoio de um supervisor para esclarecimento de dúvidas durante o trabalho de campo. O grupo coordenador nacional manteve, ao longo de todo o período de recolha de dados, uma estreita articulação com os supervisores dos entrevistadores estando sempre disponíveis para o esclarecimento de dúvidas e apoio técnico na utilização da plataforma EQ-VT.

Não obstante estar garantido o anonimato do participante durante a recolha de dados com recurso ao EQ-VT, estabeleceu-se que, para efeitos de controlo de qualidade das entrevistas, que incluiu a reinquirição parcial das mesmas, as equipas supervisionadas pela empresa de sondagens deveriam recolher o contacto e freguesia de residência de todos os participantes com o intuito da empresa proceder a uma reinquirição de uma subamostra aleatória dos indivíduos, relativamente a questões previamente definidas pelo grupo de investigadores. De notar que o grupo coordenador nacional nunca teve acesso a essa informação pessoal dos participantes, tendo apenas acesso às suas respostas. Por outro lado, a equipa de entrevistadores supervisionada directamente pelos investigadores apenas deveria recolher informação relativamente à freguesia de residência dos indivíduos entrevistados, não havendo lugar a qualquer tipo de contacto posterior com estes participantes. Esta diferença de procedimentos deveu-se, mais uma vez, à necessidade de garantir o completo anonimato dos indivíduos entrevistados.

Considerando que os entrevistadores recolheram, em cada entrevista realizada, a informação relativa à freguesia de residência dos participantes, tornou-se possível classificar o local de residência dos mesmos. Para tal, recorreu-se à Tipologia de Áreas Urbanas desenvolvida pelo Instituto Nacional de Estatística e revista em 2014 (designada por TIPAU 2014). Esta classificação obedece a uma desagregação tripartida das freguesias do território nacional em áreas predominantemente urbanas (APU), áreas mediantemente urbanas (AMU) e áreas predominantemente rurais (APR). As APU correspondem a freguesias que cumprem, pelo menos, um dos seguintes requisitos: i) o maior valor da média entre o peso da população residente na população total da freguesia e o peso da área na área total da freguesia corresponde a espaço urbano, sendo que o peso da área em espaço de ocupação predominantemente rural não ultrapassa 50% da área total da freguesia; ii) a freguesia integra a sede da Câmara Municipal e tem uma população residente superior a 5.000 habitantes; iii) e a freguesia integra total ou parcialmente um lugar com população residente igual ou superior a 5.000 habitantes, sendo que o peso da população do lugar no total da população residente na freguesia ou no total da população residente no lugar, é igual ou superior a 50%. Por sua vez, as AMU definem freguesias que cumprem, pelo menos, um dos seguintes requisitos: i) o maior valor da média entre o peso da população residente na população total da freguesia e o peso da área na área total da freguesia corresponde a Espaço Urbano, sendo que o peso da área de espaço de ocupação predominantemente rural ultrapassa 50% da área total da freguesia; ii) o maior valor da média entre o peso da população residente na população total da freguesia e o peso da área na área total da freguesia corresponde a espaço urbano em conjunto com espaço semiurbano, sendo que o peso da área de espaço de ocupação predominantemente rural não ultrapassa 50% da área total da freguesia; iii) a freguesia integra a sede da Câmara Municipal e tem uma população residente igual ou inferior a 5.000 habitantes; iv) e a freguesia integra total ou parcialmente um lugar com população residente igual ou superior a 2.000 habitantes e inferior a 5.000 habitantes, sendo que o peso da população do lugar no total da população residente na freguesia ou no total da população residente no lugar, é igual ou superior a 50%. Por último, as APR englobam freguesias que não recaem nas classificações anteriores, i.e., não são consideradas APU nem AMU.

Para o período de recolha de dados foram definidos momentos específicos para avaliação intercalar por parte dos investigadores nacionais e pela equipa de suporte do EQ-VT. No que concerne à periodicidade da avaliação sob a responsabilidade dos investigadores nacionais, esta deveria, em princípio, ocorrer diariamente. Outra periodicidade poderia ser definida em função do volume de dados recolhido pelos entrevistadores. Este acompanhamento intensivo deve-se à necessidade de avaliar cada entrevistador de 10 em 10 entrevistas, a fim de analisar o cumprimento dos critérios principais de qualidade e corrigir atempadamente situações desviantes. À semelhança do previsto no teste piloto, caso algum entrevistador tivesse quatro ou mais entrevistas assinaladas como não cumpridoras de critérios principais de qualidade, em 10 entrevistas realizadas (40% ou superior), o seu trabalho de campo era suspenso, sendo promovida uma reunião individual entre o entrevistador e os investigadores nacionais com vista a perceber o motivo do não cumprimento e apoiar a melhoria do seu desempenho. Independentemente do resultado deste apoio adicional ou retreino, o conjunto das 10 entrevistas associadas ao entrevistador era considerado comprometido e seria removido da base de dados. Se a situação de incumprimento persistisse, o grupo coordenador nacional forçava o entrevistador a abandonar o trabalho de campo. A avaliação efectuada pelo grupo coordenador nacional consistiu na elaboração de relatórios, com recurso à ferramenta QC Tool, explicação do respectivo conteúdo a cada entrevistador e apoio à melhoria do seu desempenho. Paralelamente com este processo de avaliação local e de acordo com o código de conduta, definiu-se que a equipa de suporte do EQ-VT procederia à elaboração de relatórios de avaliação para apoio à equipa nacional, com periodicidade quinzenal, e promoveria reuniões de acompanhamento com o grupo coordenador nacional, através de sessões remotas, em etapas pré-definidas da recolha de dados: 200, 400 e 600 entrevistas realizadas.

Para o período de recolha de dados, foi assumido por parte do grupo coordenador nacional o papel de suporte necessário aos entrevistadores, relativamente a questões do foro informático e a eventuais erros dessa natureza com a utilização da plataforma EQ-VT. Nos casos em que a equipa nacional não conseguisse solucionar os problemas identificados pelos entrevistadores em campo, articular-se-iam sessões remotas com um dos elementos

da equipa de suporte do EQ-VT em Roterdão, Holanda, e os entrevistadores afectados mediadas pelos investigadores nacionais.

5.3.5. Preceitos ético-legais

Considerando que este estudo utilizou cenários hipotéticos relativos a estados de saúde, não se destinando a indivíduos necessariamente doentes mas sim à população em geral e na medida em que se utilizou um instrumento de medição em saúde para obter uma valoração societária, os entrevistadores antes do início de cada entrevista disponibilizaram uma explicação cabal do objectivo da mesma e questionaram todos os indivíduos relativamente à sua vontade em participar no estudo, consubstanciando-se, nos casos positivos, num consentimento verbal, não escrito.

Previamente à implementação deste estudo, o grupo coordenador nacional notificou a Comissão Nacional de Protecção de Dados (CNPd) sobre a intenção de realizar o estudo, em Portugal, para a estimação do sistema de valores gerados pelo sistema descritivo do EQ-5D-5L. Para além de apresentar os objectivos do estudo à CNPD, o grupo coordenador nacional descreveu a metodologia a utilizar para a obtenção dos dados, definiu a população alvo e tipificou as questões envolvidas no processo de entrevista, garantindo em todo o momento o anonimato dos participantes.

Este processo de notificação culminou na autorização da CNPD para a prossecução do estudo.

5.4. Instrumento

Neste trabalho foi seguido o protocolo de entrevistas padronizado pelo grupo EuroQol para os estudos de valoração com recurso ao EQ-5D-5L – protocolo EQ-VT.

Não obstante as entrevistas serem conduzidas e realizadas por computador na aplicação EQ-VT, o que por si só promove a execução dos mesmos procedimentos por todos os entrevistadores, cada um deles dispôs ainda de um guião em papel, elaborado pelo grupo coordenador nacional, com uma explicação detalhada do conteúdo da

entrevista e com orientações obrigatórias a seguir no seu decurso. Este guião encontra-se disponível para consulta no Apêndice I.

Assim que o participante autorizava o entrevistador a realizar a entrevista, passava a controlar a utilização da plataforma EQ-VT no computador com a supervisão do entrevistador durante toda a entrevista. O controlo da aplicação apenas não era efectuado pelo participante na parte inicial da entrevista e na componente de explicação da valoração dos estados de saúde de treino.

Após efectuar o acesso à plataforma EQ-VT para registo do participante e preparação do sistema para a realização do questionário principal e do questionário local, este último definido pelos investigadores nacionais (Figuras 17 e 18), o participante era conduzido para a primeira parte do questionário principal.

Figura 17 - Acesso ao EQ-VT

Bem-vindo(a) ao estudo de valoração EQ-5D-5L EQ-5D-5L

Inicie sessão

Idioma **Nome do utilizador** **Palavra-passe**

Português (Portugal) **Iniciar sessão**

THE OFFLINE VERSION
THE EQ-5D-5L
VALUATION
STUDY
EUROQOL 2012

[Privacidade](#) | Copyright © the Euroqol Group, 2011

Figura 18 - Carregamento de questionários após registo do participante

A minha lista de tarefas EVSRM - Sistema de Gestão da Investigação do Estudo de Valoração Euroqol -

Terminar sessão

Nesta secção, é possível ver as tarefas pendentes.

Investigador a preencher os questionários dos participantes.

Portuguese valuation study - Offline Hide

*Portugal valuation study Offline**

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal Demo 2015(ned)

O questionário de valoração principal EQ-5D de 5 níveis sobre AA AA (99999) **Realizar**

Questionário adicional local sobre -offline sobre AA AA (99999)Este questionário destina-se a residentes em Portugal. **Saltar** **Realizar**

O protocolo de valoração EQ-VT consiste em cinco partes distintas: a) boas-vindas e apresentação do estudo; b) introdução; c) valoração cTTO; d) valoração DCE e e) agradecimentos. Seguidamente, apresenta-se uma explicação detalhada de cada uma delas.

a) Boas-vindas e apresentação do estudo

Nesta fase inicial, o entrevistador dava a boas-vindas ao participante e agradecia a sua participação no estudo (Figura 19). Fazia um pequeno enquadramento do estudo, explicava ainda os seus objectivos e apresentava a estrutura da entrevista bem como a duração prevista (entre 30 a 45 minutos).

Figura 19 - Boas-vindas e apresentação do estudo

Participante aleatorizado com sucesso para o(s) seguinte(s) braço(s) do estudo: TTO Group 7, LEFTTORIGHT, DCE Group 18

15513327 -

Terminar sessão

PERSPETIVAS SOBRE A SAÚDE

Obrigado por concordar em participar neste estudo. Estamos a realizar esta entrevista para permitir melhoramentos no sistema de saúde. Pede-se a elementos da população em geral, na qual está incluído(a), que deem o seu contributo para o estudo. As perguntas referem-se às diferentes opiniões das pessoas sobre a saúde e a doença. Ser-lhe-á pedido que pense sobre alguns estados de saúde por que pode ou não ter passado na sua vida.

Não há respostas certas ou erradas. Diga apenas o que pensa. Com base na experiência anterior, o estudo deverá demorar entre 30 e 45 minutos. Caso não se sinta à vontade para responder às perguntas, pode interromper em qualquer momento. As suas respostas serão tratadas com rigorosa confidencialidade.

Está preparado(a) para começar?

← Anterior

Seguinte →

Idioma atual

Português (Portugal)

Linha de apoio do estudo

+1 (866) 338-4894

Sessão iniciada como:

INV_PORT_DEMO

ID de grupo:

Portugal Demo 2015(ned)

b) Introdução

De seguida, o participante era convidado a responder a algumas perguntas consideradas de contexto. Em primeiro lugar, o entrevistador perguntava como se sentia o participante no dia da entrevista, relativamente ao seu estado de saúde, seguindo ordenadamente as cinco dimensões do EQ-5D: mobilidade, cuidados pessoais, actividades habituais, dor/mal-estar e ansiedade/depressão (Figuras 20 a 24).

Figura 20 - EQ-5D-5L: avaliação actual da mobilidade

15513327 -

Seleccione UM quadrado que melhor descreva a sua saúde HOJE, clicando nele.

MOBILIDADE

Não tenho problemas em andar	<input type="checkbox"/>
Tenho problemas ligeiros em andar	<input type="checkbox"/>
Tenho problemas moderados em andar	<input type="checkbox"/>
Tenho problemas graves em andar	<input type="checkbox"/>
Sou incapaz de andar	<input type="checkbox"/>

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

 *Sessão iniciada como:*
INV_PORT_DEMO

 *ID de grupo:*
Portugal
Demo
2015(ned)

← Anterior

Seguinte →

Figura 21 - EQ-5D-5L: avaliação actual dos cuidados pessoais

15513327 -

Seleccione UM quadrado que melhor descreva a sua saúde HOJE, clicando nele.

CUIDADOS PESSOAIS

Não tenho problemas em me lavar ou vestir	<input type="checkbox"/>
Tenho problemas ligeiros em me lavar ou vestir	<input type="checkbox"/>
Tenho problemas moderados em me lavar ou vestir	<input type="checkbox"/>
Tenho problemas graves em me lavar ou vestir	<input type="checkbox"/>
Sou incapaz de me lavar ou vestir sozinho(a)	<input type="checkbox"/>

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

 *Sessão iniciada como:*
INV_PORT_DEMO

 *ID de grupo:*
Portugal
Demo
2015(ned)

← Anterior

Seguinte →

Figura 22 - EQ-5D-5L: avaliação actual das actividades habituais

15513327 -

Seleccione UM quadrado que melhor descreva a sua saúde HOJE, clicando nele.

ATIVIDADES HABITUAIS (ex. trabalho, estudos, atividades domésticas, atividades em família ou de lazer)

Não tenho problemas em desempenhar as minhas atividades habituais	<input type="checkbox"/>
Tenho problemas ligeiros em desempenhar as minhas atividades habituais	<input type="checkbox"/>
Tenho problemas moderados em desempenhar as minhas atividades habituais	<input type="checkbox"/>
Tenho problemas graves em desempenhar as minhas atividades habituais	<input type="checkbox"/>
Sou incapaz de desempenhar as minhas atividades habituais	<input type="checkbox"/>

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

 *Sessão iniciada como:*
INV_PORT_DEMO

 *ID de grupo:*
Portugal
Demo
2015(ned)

Anterior ←**Seguinte** →

Figura 23 - EQ-5D-5L: avaliação actual da dor/mal-estar

15513327 -

Seleccione UM quadrado que melhor descreva a sua saúde HOJE, clicando nele.

DOR/MAL-ESTAR

Não tenho dores ou mal-estar	<input type="checkbox"/>
Tenho dores ou mal-estar ligeiros	<input type="checkbox"/>
Tenho dores ou mal-estar moderados	<input type="checkbox"/>
Tenho dores ou mal-estar graves	<input type="checkbox"/>
Tenho dores ou mal-estar extremos	<input type="checkbox"/>

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

 *Sessão iniciada como:*
INV_PORT_DEMO

 *ID de grupo:*
Portugal
Demo
2015(ned)

Anterior ←**Seguinte** →

Figura 24 - EQ-5D-5L: avaliação actual da ansiedade/depressão

15513327 -

Seleccione UM quadrado que melhor descreva a sua saúde HOJE, clicando nele.

ANSIEDADE/DEPRESSÃO

Não estou ansioso(a) ou deprimido(a)

Estou ligeiramente ansioso(a) ou deprimido(a)

Estou moderadamente ansioso(a) ou deprimido(a)

Estou gravemente ansioso(a) ou deprimido(a)

Estou extremamente ansioso(a) ou deprimido(a)

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal
Demo
2015(ned)

Anterior ←
Seguinte →

Na fase seguinte, pedia-se ao participante para indicar numa escala numerada de 0 a 100 (VAS) qual o valor que atribuía ao seu estado de saúde nesse momento (Figura 25).

Figura 25 - VAS

15513327 -

- Gostaríamos de saber o quanto a sua saúde está boa ou má HOJE.
- A escala está numerada de 0 a 100.
- 100 significa a melhor saúde que possa imaginar.
- 0 significa a pior saúde que possa imaginar.
- Por favor, clique na escala para indicar como a sua saúde está HOJE.

A SUA SAÚDE HOJE

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal
Demo
2015(ned)

A melhor saúde que possa imaginar

100

90

80

70

60

50

40

30

20

10

0

A pior saúde que possa imaginar

Anterior ←
Seguinte →

E, por fim, eram efectuadas algumas perguntas pessoais relativamente à idade, género e experiência prévia com alguma doença grave (Figura 26).

Figura 26 - Perguntas de contexto pessoal

15513327 -

Perguntas sobre dados pessoais

1. Já teve contacto com alguma doença grave?

sua Sim não

de algum familiar Sim não

enquanto cuidador(a) Sim não

2. Qual é a sua idade? anos

3. É do sexo masculino ou feminino? Masculino Feminino

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal Demo 2015(ned)

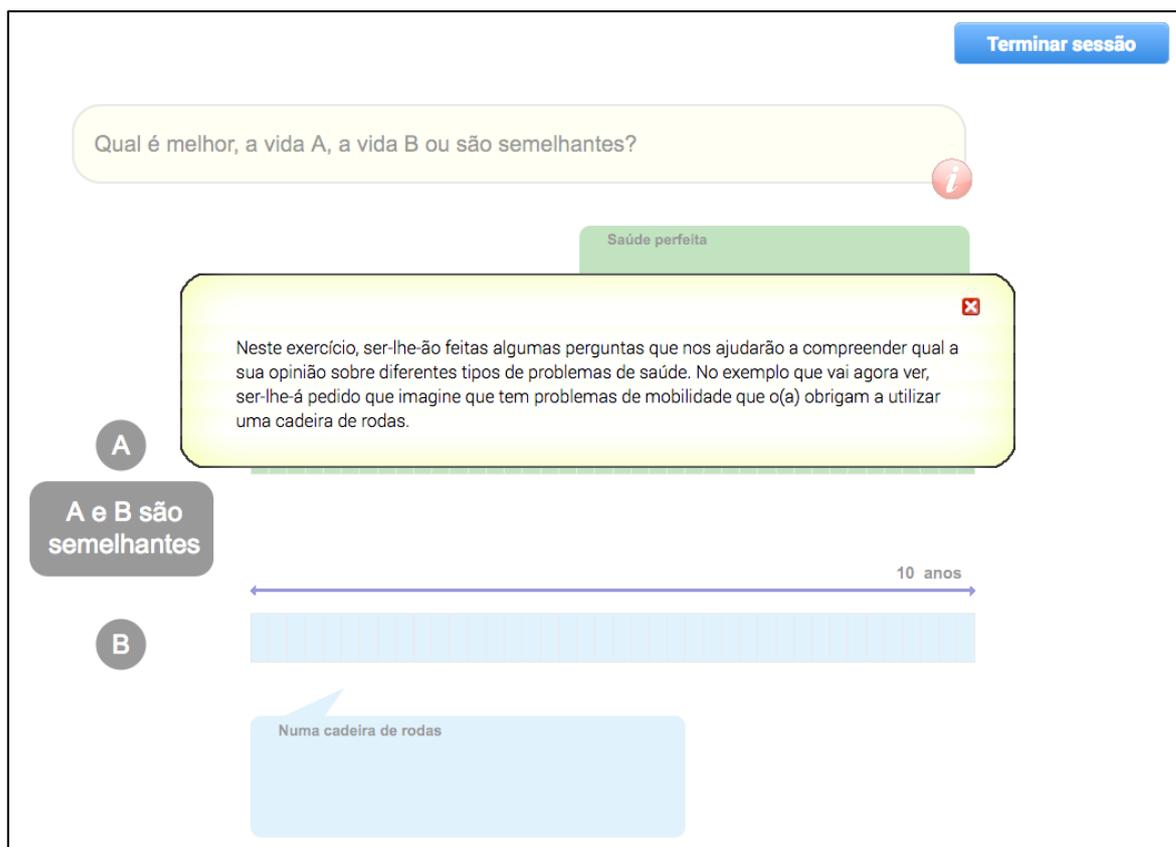
Anterior ←

Seguinte →

c) Valoração cTTO

Antes de se iniciar a primeira técnica de valoração – o TTO composto ou cTTO – o participante passava primeiro por uma parte de treino constituída por duas fases (Figura 27). O objectivo deste treino era introduzir o indivíduo na técnica cTTO e facilitar a valoração dos estados de saúde melhores que morte e dos estados de saúde considerados piores que morte. A primeira fase do treino era denominada de “exemplo da cadeira de rodas”.

Figura 27 - Introdução ao exemplo da cadeira de rodas



Neste exemplo, seguindo o Protocolo EQ-VT, pedia-se ao participante para comparar 10 anos em saúde perfeita (Vida A) com 10 anos num estado de saúde relativamente ligeiro mas que o obrigava a viver numa cadeira de rodas (Vida B).¹⁰ Note-se que a duração associada à permanência num estado de saúde definida para os estudos de valoração do EQ-5D e à técnica cTTO é de 10 anos (t). Dado tratar-se de um estado de saúde ligeiro, ainda que o indivíduo tenha que permanecer numa cadeira de rodas, considera-se que se trata de um estado de saúde melhor que morte. Era explicado ao participante que, durante esses 10 anos, não havia qualquer tratamento que modificasse a sua situação e que, findo esse período, quer estivesse no estado de saúde perfeita (Vida A) quer estivesse no estado de saúde ligeiro na cadeira de rodas (Vida B), morreria (Figura 28). O objectivo deste processo iterativo é que o participante esteja disposto a sacrificar anos de vida no estado de saúde perfeita para evitar permanecer no estado de saúde ligeiro

¹⁰ O indivíduo dará a interpretação que desejar para os termos “relativamente ligeiro” e “saúde perfeita”.

na cadeira de rodas, até alcançar o seu ponto de indiferença em que considera que o número de anos em saúde perfeita (x) é semelhante ao número de anos no estado de saúde ligeiro na cadeira de rodas (t), em que $x < t$ (Figura 29). Atingindo o ponto de indiferença, o valor atribuído ao estado de saúde será de $\frac{x}{t}$ ou, nesse caso, $\frac{x}{10}$.

Figura 28 - Exemplo da cadeira de rodas (início da iteração): estado de saúde melhor que morte (TTO convencional)

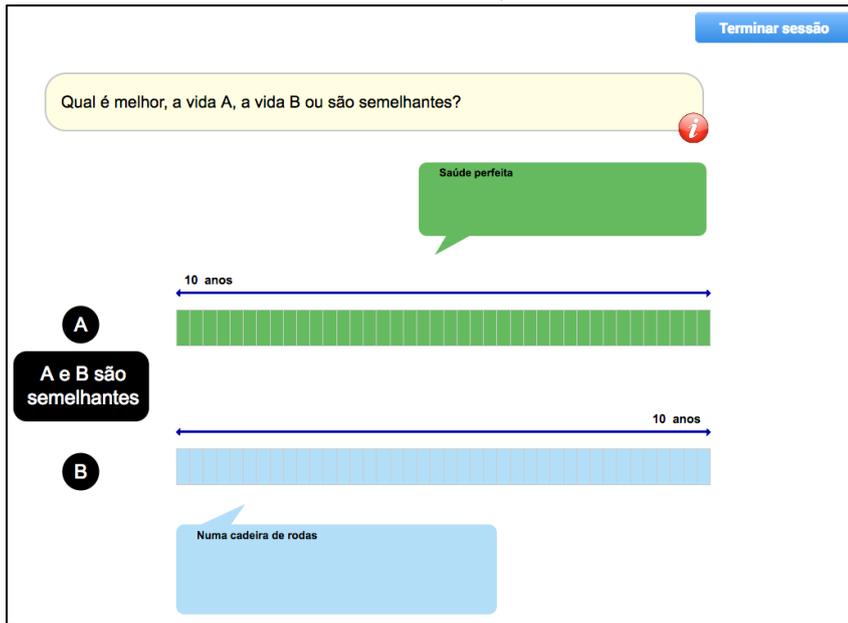
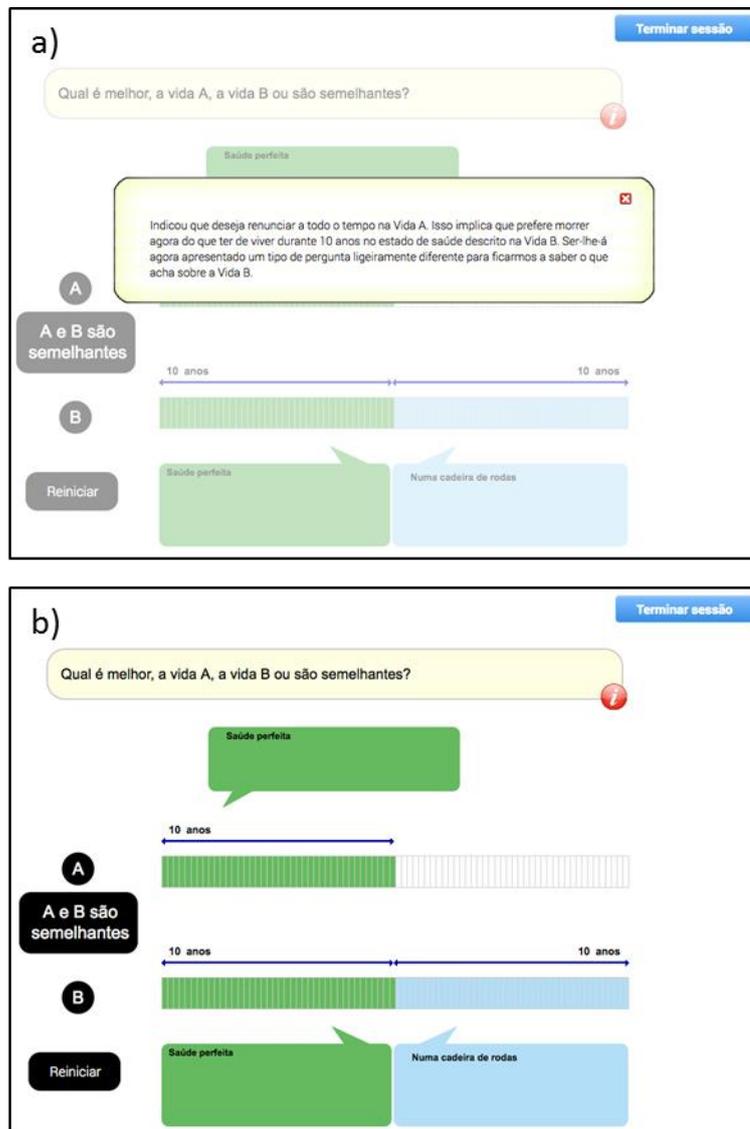


Figura 29 - Exemplo da cadeira de rodas (final da iteração): estado de saúde melhor que morte (TTO convencional)

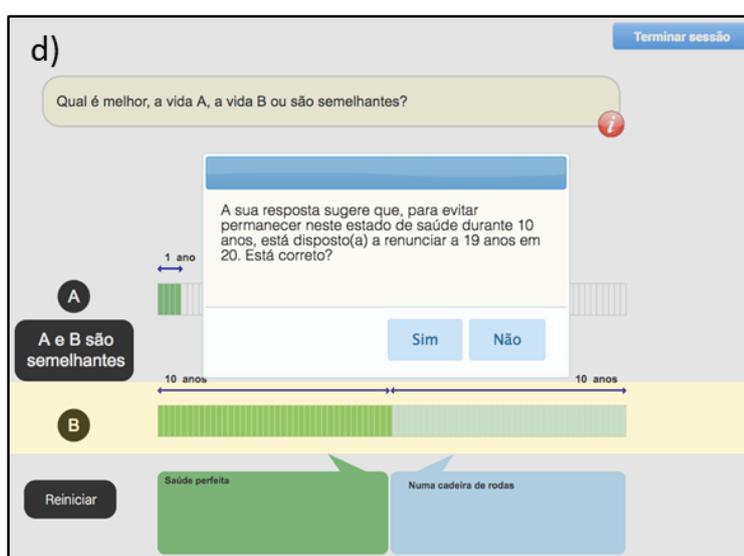
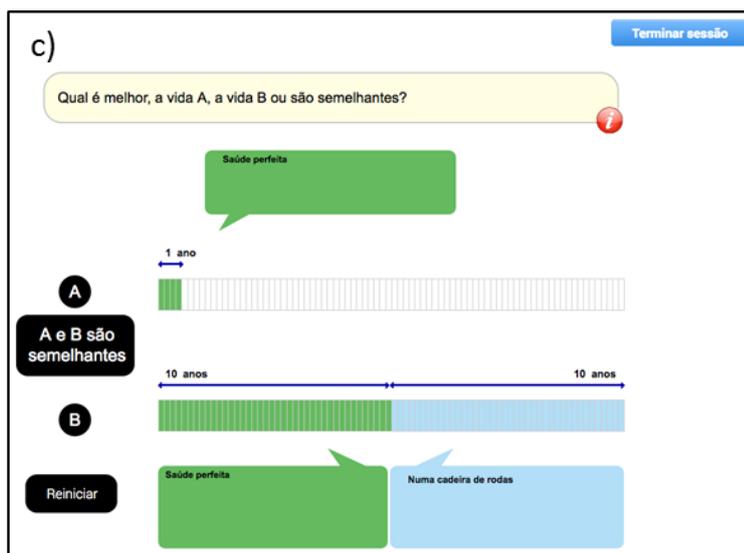


Para finalizar o exemplo da cadeira de rodas, pedia-se de seguida ao participante que imaginasse que, estando à mesma em cadeira de rodas, o estado de saúde descrito na Vida B era tão mau que preferiria morrer naquele momento do que ter que viver nessa situação, durante 10 anos. Com esta variação ao exemplo da cadeira de rodas pretendia-se introduzir o participante na valoração de estados de saúde considerados piores que morte (Figura 30).

Figura 30 - Exemplo da cadeira de rodas: estado de saúde considerado pior que morte¹¹



¹¹ A sucessão de imagens é meramente representativa dado que depende das respostas dos participantes.

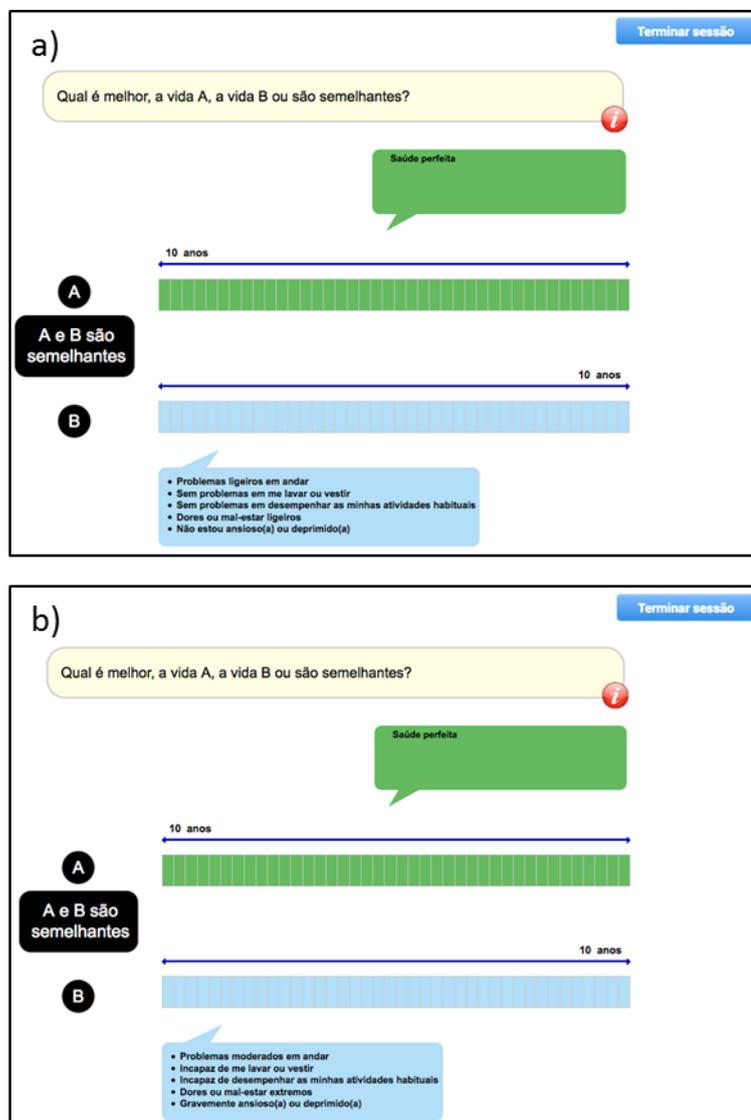


Tendo em conta que os indivíduos terão opiniões e crenças distintas relativamente a permanecerem vivos em circunstâncias de saúde miserável e com o intuito de respeitar as suas opiniões, pedia-se apenas que o participante imaginasse que preferia morrer para efeitos de compreensão do exemplo de treino, salvaguardando que não se tratava de uma verdadeira valoração. Essa ocorrência num momento posterior da entrevista nas 10 perguntas de valoração. O que este exemplo faz, tal como explicado anteriormente, é permitir a eliciação de valores negativos, adicionando 10 anos em saúde perfeita antes de se alcançar o estado de saúde a ser valorado na Vida B que durará também 10 anos (*lead-time* TTO). Uma vez mais pedia-se ao participante que sacrificasse anos na Vida A (x) para evitar permanecer 20 anos na Vida B. Nestas circunstâncias, o valor obtido para o estado

de saúde é dado por $(x - 10)/10$. Este rácio implica que o menor valor que se pode obter é -1 (quando $x = 0$). Neste processo de troca de tempo, a menor unidade de tempo transaccionável equivalia a seis meses ou 0,05.

Findo o exemplo da cadeira de rodas, apresentava-se ao participante mais uma tarefa de treino que consistia na valoração cTTO de três estados de saúde descritos seguindo o sistema descritivo do EQ-5D. Em todas as entrevistas estes estados de saúde tinham a mesma descrição, i.e. estavam fixos, e representavam sequencialmente um estado de saúde com alterações ligeiras, graves e moderadas (Figura 31).

Figura 31 - Valoração cTTO: estados de saúde de treino



c)

Qual é melhor, a vida A, a vida B ou são semelhantes?

Terminar sessão

Saúde perfeita

10 anos

A

A e B são semelhantes

10 anos

B

- Sem problemas em andar
- Incapaz de me lavar ou vestir
- Problemas graves em desempenhar as minhas atividades habituais
- Sem dores ou mal-estar
- Não estou ansioso(a) ou deprimido(a)

d)

Qual é melhor, a vida A, a vida B ou são semelhantes?

Terminar sessão

Saúde perfeita

Chegou ao fim da secção de treino. Ser-lhe-ão agora feitas 10 perguntas. Agora, ser-lhe-á apresentado um novo estado de saúde. Observe-o atentamente na caixa azul para a Vida B. Qual é melhor, a vida A, a vida B ou são semelhantes?

A

A e B são semelhantes

10 anos

B

- Incapaz de andar
- Problemas moderados em me lavar ou vestir
- Problemas graves em desempenhar as minhas atividades habituais
- Sem dores ou mal-estar
- Ligeiramente ansioso(a) ou deprimido(a)

Terminadas as tarefas de treino, cada participante era convidado a valorar, de acordo com as suas preferências, 10 estados de saúde através da técnica cTTO. Os 10 estados de saúde eram aleatoriamente atribuídos pelo EQ-VT, de acordo com o desenho pré-definido do estudo e explicitado em subcapítulo posterior. No entanto, em cada conjunto de 10 estados de saúde era sempre incluído o estado de saúde com a pior descrição possível, i.e., com problemas extremos em todas as cinco dimensões. Considerando que cada dimensão ocupa sempre a mesma ordem sequencial (mobilidade → cuidados pessoais → actividades habituais → dor/mal-estar → ansiedade/depressão) e que o nível 1 representa a ausência de problemas nessa dimensão, o nível 2 constitui a

presença de problemas ligeiros, o nível 3 representa problemas moderados, o nível 4 problemas graves e o nível 5 problemas extremos, ao pior estado de saúde descrito pelo EQ-5D-5L é atribuída a designação de 55555 que representa um estado de saúde em que o indivíduo se encontra incapaz de andar, incapaz de realizar os seus cuidados pessoais (lavar-se e vestir-se sozinho), incapaz de desempenhar as suas actividades habituais, apresenta dor/mal-estar extremos e ansiedade/depressão extremas. A Figura 32 representa um exemplo real de valoração de um estado de saúde descrito pelo EQ-5D-5L com recurso ao cTTO.

Figura 32 - Valoração de estado de saúde com recurso ao cTTO

Qual é melhor, a vida A, a vida B ou são semelhantes?

Terminar sessão

Saúde perfeita

10 anos

A

A e B são semelhantes

B

10 anos

- Incapaz de andar
- Problemas moderados em me lavar ou vestir
- Problemas graves em desempenhar as minhas atividades habituais
- Sem dores ou mal-estar
- Ligeiramente ansioso(a) ou deprimido(a)

Após a valoração dos 10 estados de saúde, o participante era convidado a avaliar o grau de dificuldade da tarefa, através de uma escala de cinco pontos de Likert, no que respeita às questões colocadas, à distinção entre os estados de saúde descritos e o ponto de indiferença entre as vidas apresentadas (Figura 33).

Figura 33 - Avaliação da dificuldade da tarefa cTTO

Por favor, diga-nos o que pensou das perguntas a que acabou de responder, em que comparou duas “vidas” diferentes.

	Concordo			Não concordo	
Achei fácil perceber as questões que me foram colocadas.	<input type="radio"/>				
Achei fácil distinguir entre as vidas que me pediram para ponderar.	<input type="radio"/>				
Achei difícil decidir quais os pontos exatos em que a Vida A e a Vida B eram semelhantes.	<input type="radio"/>				

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

 *Sessão iniciada como:*
INV_PORT_DEMO

 *ID de grupo:*
Portugal Demo 2015(ned)

Seguinte 

De seguida, os 10 estados de saúde anteriormente valorados eram apresentados, ordenados de acordo com as respostas dadas pelo participante. Na parte superior do ecrã apresentavam-se os estados de saúde considerados melhores pelo participante e na parte inferior do ecrã, os considerados piores. Os estados de saúde para os quais o participante havia atribuído o mesmo valor surgiram a par na mesma linha. De notar que sempre que o participante considerava um estado de saúde como sendo pior que morte este surgia no ecrã assinalado a azul escuro. O objectivo desta etapa não era corrigir qualquer avaliação efectuada pelo participante mas sim permitir a sua revisão e assinalar eventuais incoerências, a fim de poderem ser removidas da base de dados aquando da modelação econométrica. Assim, nesta fase designada por módulo de feedback, pedia-se ao entrevistado que revisse a ordem das respostas dadas e assinalasse algum estado que eventualmente considerasse mal posicionado (Figura 34).

Figura 34 - Módulo de feedback

Melhores estados de saúde na parte superior

Terminar sessão

- Sem problemas em andar
- Sem problemas em me lavar ou vestir
- Sem problemas em desempenhar as minhas atividades habituais
- Sem dores ou mal-estar
- Ligeiramente ansioso(a) ou deprimido(a)

- Problemas ligeiros em andar
- Sem problemas em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar moderados
- Gravemente ansioso(a) ou deprimido(a)

- Sem problemas em andar
- Problemas ligeiros em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar moderados
- Gravemente ansioso(a) ou deprimido(a)

- Problemas moderados em andar
- Problemas ligeiros em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Sem dores ou mal-estar
- Gravemente ansioso(a) ou deprimido(a)

- Problemas moderados em andar
- Problemas moderados em me lavar ou vestir
- Problemas ligeiros em desempenhar as minhas atividades habituais
- Dores ou mal-estar extremos
- Moderadamente ansioso(a) ou deprimido(a)

- Problemas ligeiros em andar
- Problemas moderados em me lavar ou vestir
- Problemas ligeiros em desempenhar as minhas atividades habituais
- Dores ou mal-estar graves
- Ligeiramente ansioso(a) ou deprimido(a)

- Problemas ligeiros em andar
- Problemas graves em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar graves
- Ligeiramente ansioso(a) ou deprimido(a)

- Incapaz de andar
- Problemas moderados em me lavar ou vestir
- Problemas graves em desempenhar as minhas atividades habituais
- Sem dores ou mal-estar
- Ligeiramente ansioso(a) ou deprimido(a)

- Incapaz de andar
- Incapaz de me lavar ou vestir
- Problemas ligeiros em desempenhar as minhas atividades habituais
- Dores ou mal-estar ligeiros
- Extremamente ansioso(a) ou deprimido(a)

- Incapaz de andar
- Incapaz de me lavar ou vestir
- Incapaz de desempenhar as minhas atividades habituais
- Dores ou mal-estar extremos
- Extremamente ansioso(a) ou deprimido(a)

Piores estados de saúde na parte inferior

Seguinte

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal Demo 2015(ned)

Por fim, convidava-se o participante a introduzir na plataforma algum comentário que considerasse pertinente relativamente à ordenação dos estados de saúde no módulo de feedback. Caso verificasse a necessidade de consultar novamente a ordem de classificação dos estados de saúde, a plataforma permitia aceder à etapa anterior (Figura 35).

Figura 35 - Caixa de comentários associada à tarefa cTTO

15513327 -

Utilize a caixa de texto abaixo para introduzir quaisquer comentários relativamente à ordem de classificação dos estados de saúde indicados no ecrã anterior. Se quiser consultar novamente a ordem de classificação dos estados de saúde, clique no botão "Mostrar ordenação".

← Anterior

Seguinte →

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal
Demo
2015(ned)

d) Valoração DCE

Terminada a primeira tarefa de valoração, o EQ-VT prosseguia para a segunda e última tarefa agora com recurso à técnica DCE (Figura 36).

Figura 36 - Apresentação da tarefa de valoração DCE

15513327 -

Em cada uma das perguntas que se seguem serão apresentados dois estados de saúde diferentes. Observe-os atentamente. Escolha aquele que, na sua opinião, é melhor premindo o botão "A" ou "B".

- Problemas moderados em andar
- Problemas graves em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar ligeiros
- Gravemente ansioso(a) ou deprimido(a)

A

- Problemas graves em andar
- Problemas moderados em me lavar ou vestir
- Problemas ligeiros em desempenhar as minhas atividades habituais
- Dores ou mal-estar graves
- Extremamente ansioso(a) ou deprimido(a)

B

Terminar sessão

Idioma atual
Português (Portugal)

Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal
Demo
2015(ned)

Esta tarefa consistia na apresentação de sete pares de estados de saúde descritos de acordo com as cinco dimensões e níveis constantes do EQ-5D-5L. Neste caso, não estava implícita qualquer duração ou tempo de vida associado à tarefa de valoração. Apenas se pedia ao participante que escolhesse entre a Vida A ou a Vida B, considerando que representavam estados de saúde distintos. A Figura 37 apresenta um exemplo de um par de estados de saúde para valoração com recurso à DCE.

Figura 37 - Valoração de um par de estados de saúde com recurso à DCE

The figure displays two screenshots of a web-based interface for Discrete Choice Experiment (DCE) valuation. Both screenshots show a central question: "Qual é melhor, o estado 'A' ou o estado 'B'?" (Which is better, state 'A' or state 'B'?).

State A (Left Panel):

- Problemas moderados em andar
- Problemas graves em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar ligeiros
- Gravemente ansioso(a) ou deprimido(a)

State B (Right Panel):

- Problemas graves em andar
- Problemas moderados em me lavar ou vestir
- Problemas ligeiros em desempenhar as minhas atividades habituais
- Dores ou mal-estar graves
- Extremamente ansioso(a) ou deprimido(a)

Right Sidebar (Common to both):

- Terminar sessão
- Idioma atual: Português (Portugal)
- Linha de apoio do estudo: +1 (856) 338-4894
- Sessão iniciada como: INV_PORT_DEMO
- ID de grupo: Portugal Demo 2015(ned)

Bottom Screenshot (Feedback Dialog):

A sua resposta sugere que prefere o estado de saúde B ao estado de saúde A. Está correto? (Esta é a pergunta 1 de 7 perguntas semelhantes)

Buttons: Sim, Não

Neste estudo, os valores dos estados de saúde com recurso à técnica DCE não foram estimados numa escala de utilidade cuja métrica se situa entre 0 (morte) e 1 (saúde perfeita). Por este motivo, os resultados obtidos com esta técnica não podem ser usados de forma independente neste estudo para estimar o valor dos estados de saúde (Rowen et al. 2015; Stolk et al. 2010).

Finda a tarefa DCE e à semelhança do que já havia ocorrido com a técnica cTTO, o participante era convidado a avaliar o grau de dificuldade da tarefa de valoração DCE, no que respeita às questões colocadas, à distinção entre os estados de saúde descritos e a decisão entre as vidas (Figura 38).

Figura 38 - Avaliação da dificuldade da tarefa DCE

15513327 -

Por favor, diga-nos o que pensou das perguntas a que acabou de responder, em que comparou dois estados de saúde diferentes.

	Concordo			Não concordo
Achei fácil perceber as questões que me foram colocadas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Achei fácil distinguir entre os estados de saúde que me pediram para ponderar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Achei difícil decidir como responder .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

← Anterior

Seguinte →

Terminar sessão

Idioma atual
Português (Portugal)

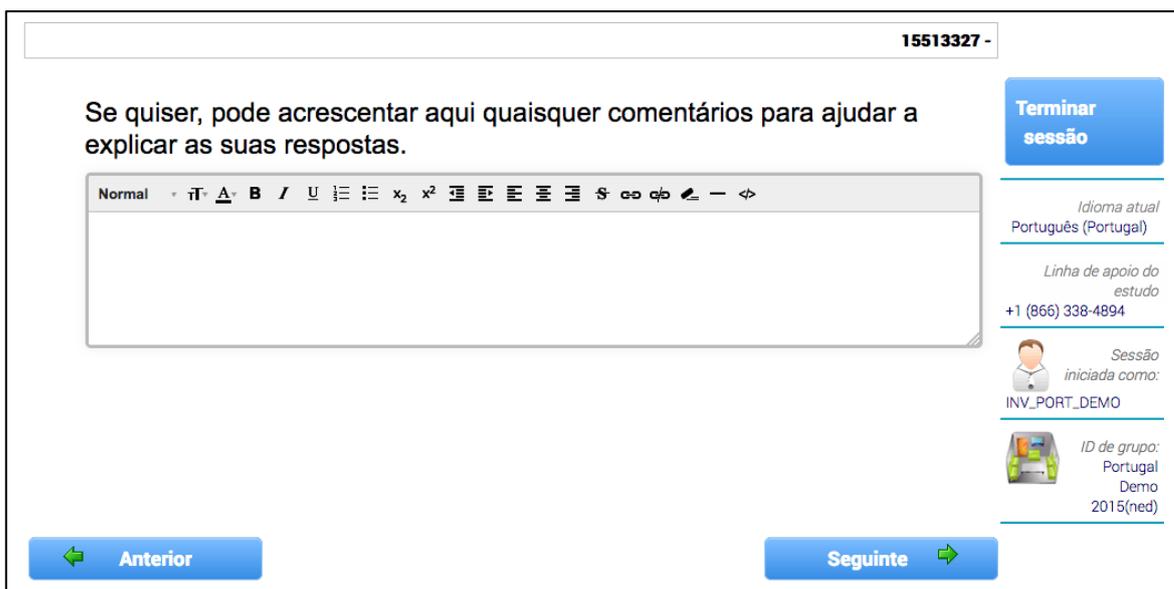
Linha de apoio do estudo
+1 (866) 338-4894

Sessão iniciada como:
INV_PORT_DEMO

ID de grupo:
Portugal Demo
2015(ned)

Adicionalmente, era convidado a introduzir algum comentário opcional que considerasse oportuno para a explicação das suas respostas nesta tarefa de valoração (Figura 39).

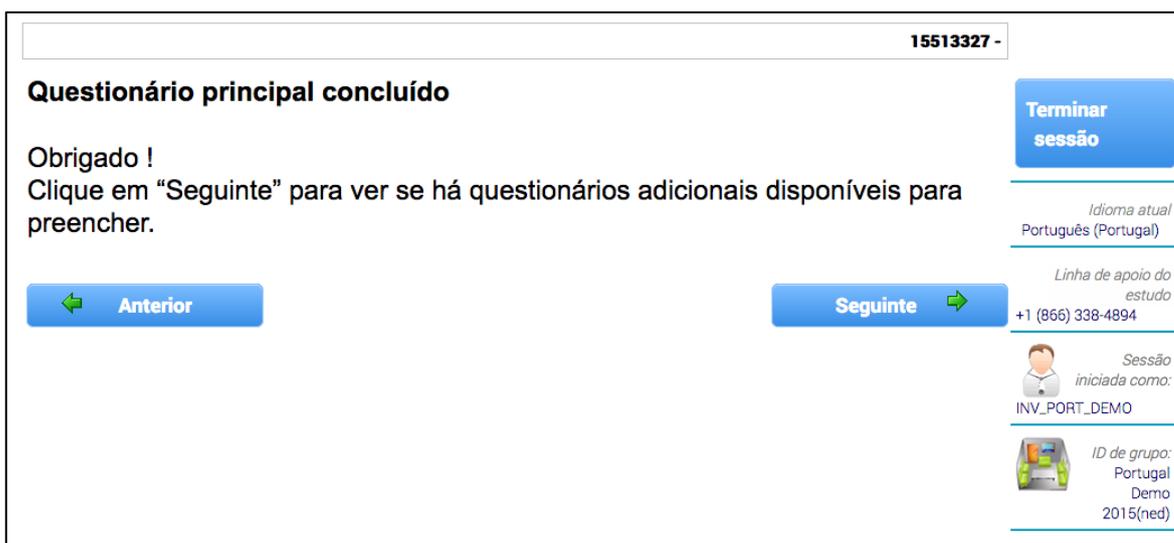
Figura 39 - Caixa de comentários associada à tarefa DCE



e) Agradecimentos

Para finalização do questionário principal, o entrevistador agradecia o esforço e a colaboração do participante (Figura 40), passando de seguida ao segundo questionário de perguntas locais definidas pela equipa nacional de investigadores.

Figura 40 - Finalização do questionário principal e agradecimentos



5.4.1. Questionário local

O questionário local, desenvolvido pelo grupo coordenador nacional e previsto no código de conduta assumido com o grupo EuroQol, consistiu num conjunto de perguntas específicas, introduzidas com vista a aprofundar a análise dos resultados obtidos com o estudo de valoração. Estas perguntas adicionais incluíram questões de cariz sociodemográfico e outras com vista a uma melhor compreensão das preferências dos participantes nas tarefas de valoração. Neste âmbito, pretendeu-se perceber quais as dimensões do EQ-5D-5L (mobilidade, cuidados pessoais, actividades habituais, dor/mal-estar e ansiedade/depressão) mais relevantes para o participante quando se aplicou a técnica de valoração cTTO e a técnica de valoração DCE, se existia experiência prévia com alguma doença, se o participante atribuía maior importância à longevidade ou à qualidade de vida, qual a sua principal preocupação em situação de doença grave e crenças perante a morte. Estas foram as variáveis que os investigadores nacionais consideraram relevantes com vista a uma análise adicional e complementar dos resultados do estudo. Embora estes resultados não sejam avaliados na totalidade nesta tese, por se afastarem do seu objectivo principal, são analisados noutra trabalho que se encontra em preparação.

5.5. Controlo de qualidade

Neste estudo foram previstas duas medidas de controlo de qualidade dos dados: uma implementada sob a supervisão do grupo EuroQol (QC Tool) e outra criada pelos investigadores nacionais (Reinquirição).

5.5.1. Ferramenta de controlo de qualidade: QC Tool

Os estudos anteriores de valoração do EQ-5D-5L evidenciaram consequências indesejadas na qualidade dos dados resultantes do efeito causado pelo desempenho dos entrevistadores (Augustovski et al. 2016; Ramos-Goñi et al. 2014).

Por esse motivo, o grupo EuroQol decidiu criar uma ferramenta de controlo de qualidade com o objectivo de monitorizar o desempenho dos entrevistadores. Esta

ferramenta – QC Tool – é um ficheiro automatizado que, quando carregado com os dados em bruto, com a informação dos relatórios parcelares cTTO e DCE e com o ficheiro de logística (cf. Figura 10), elabora um conjunto de estatísticas disponibilizadas em formato Excel, para análise dos investigadores, e um relatório em formato Word, para análise por parte dos investigadores e divulgação junto dos entrevistadores (Figura 41).

No decorrer de uma entrevista, a plataforma EQ-VT está preparada para recolher informação relativa a toques no teclado, número de movimentos executados no processo iterativo do cTTO, tempo em que ocorrem as várias fases e etapas cumpridas. Esta informação é recolhida com vista à monitorização do comportamento dos entrevistadores e da sua adesão ao protocolo da entrevista e, no limite, à validação dos dados.

Figura 41 - QC Tool

The screenshot displays the QC Tool interface with the following elements:

- Header:** "These Excel macros have been developped to help on the quality control (QC) process for EQ-5D-5L country specific Valuation Studies".
- Logo:** EQ-5D logo on the left.
- Footer:** "Powered by EuroQol Group Office Contact support: jramos@euroqol.org" and "Tool version: 2.4".
- Important note:** "This software could not run properly if you edit/remove yellow sheets manually".
- First step:** "Import data from EQ-VT reports". Instructions: "Import data from EQ-VT website reports by selecting the files and pressing 'Import and Combine Data Files'. Please, do NOT edit the EQ-VT report manually, they have to be the originals downloaded from EQ-VT site." Buttons: "Select DCE File", "Select TTO File", "Select PI File", and "Import and Combine Data Files".
- Second step:** "Clean data to drop invalid observations". Instructions: "Cleaning can be done in 3 ways (Yellow buttons). 1) 'Remove data by participants ID or Regular Expression': This button will ask about an external ID (Regular expressions as '*r99??' are allowed) 2 & 3) 'Keep data by participant ID list' and 'Remove data by participant ID list': first to click the 'create the participants list' button, after which a worksheet sheet called 'ID list' is created. In this sheet all internal IDs to be kept or removed have to be listed. (Please, analyze PI file to create the internal ID list)". Buttons: "Remove data by participants ID or Regular Expression", "Create the participants ID list", "Remove data by the participant ID list", and "Keep data by the participant ID list".
- Third step:** "Create Word report". Button: "Word report".
- Bottom right:** "Remove All" button.

Assumindo que os entrevistadores conduzem a entrevista seguindo fielmente o guião disponibilizado para o efeito, o QC Tool compara as etapas previstas no protocolo com o efectivamente sucedido numa entrevista.

Neste processo de controlo foram definidos quatro critérios principais na tarefa cTTO cujo incumprimento, a suceder individualmente ou em grupo, obriga a assinalar uma entrevista como sendo de reduzida qualidade e, portanto, potencialmente não válida:

- i) não introdução do estado de saúde pior que morte (*lead-time* TTO) no exemplo de treino da cadeira de rodas;
- ii) tempo inferior a três minutos para executar a componente de treino do exemplo da cadeira de rodas (treino dos estados melhor e pior que morte);
- iii) inconsistência flagrante na valoração dos estados de saúde no cTTO (o estado de saúde 55555 não é valorado como o pior estado e a valoração a este atribuída é pelo menos 0,5 superior à valoração do estado de saúde que obteve o pior valor);
- iv) tempo inferior a cinco minutos para a realização das 10 tarefas de valoração cTTO.

Neste contexto, o desempenho do entrevistador e a sua adesão ao protocolo foram avaliados em função do número de entrevistas assinaladas e de acordo com o seguinte procedimento:

- quando um entrevistador realizou um conjunto de 10 entrevistas, o seu desempenho e adesão ao protocolo foram avaliados com a ferramenta QC Tool;
- se quatro ou mais entrevistas em 10 realizadas foram assinaladas como sendo de reduzida qualidade, todas as entrevistas desse conjunto de 10 foram removidas da base de dados e o entrevistador foi retreinado;
- após a realização de mais 10 entrevistas pelo entrevistador, o desempenho e a adesão ao protocolo foram reavaliados. Se, de novo, quatro ou mais entrevistas do conjunto de 10 realizadas foram assinaladas como sendo de reduzida qualidade, todos os dados foram removidos e o entrevistador foi forçado a abandonar o estudo;
- ao longo de todo o estudo, o grupo coordenador nacional avaliou continuamente o desempenho dos entrevistadores e a sua adesão ao protocolo, utilizando a ferramenta QC Tool e analisando os relatórios de avaliação resultantes. A sua actuação sucedeu não só quando surgiram os

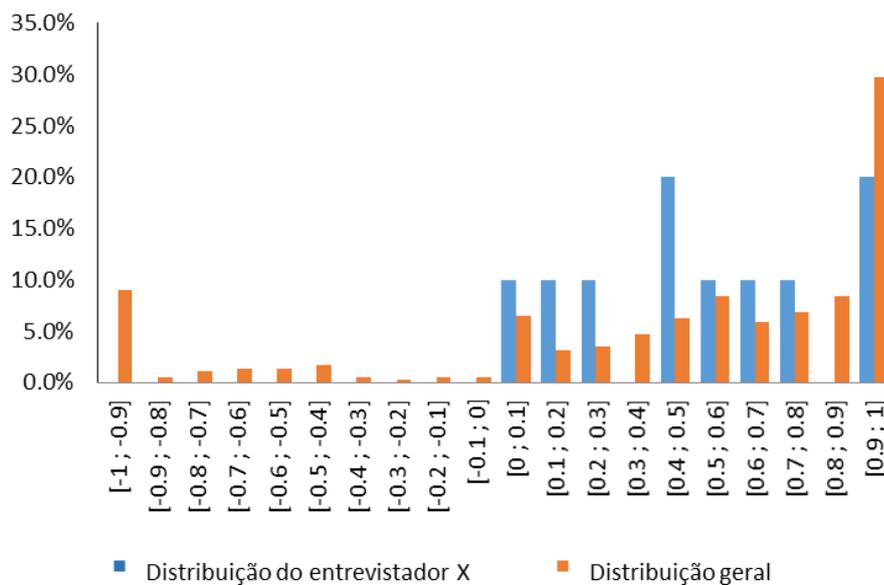
problemas descritos anteriormente mas também ao primeiro sinal de agravamento do desempenho dos entrevistadores.

Para além dos critérios principais identificados anteriormente, os dados foram periodicamente analisados tendo por base situações já identificadas em estudos anteriores de valoração do EQ-5D-5L, designadamente:

- i) existência de poucas valorações piores que morte (negativas) associadas a um entrevistador;
- ii) agrupamento de valores (picos) em torno de certos valores tais como 1; 0,5; 0; -0,5 e -1;
- iii) valoração reduzida para estados de saúde ligeiros.

A título de ilustração destes critérios, o gráfico 1 apresenta um exemplo da distribuição dos dados cTTO com identificação de um pico em torno de 0,5 e ausência de valoração de estados de saúde piores que morte (inexistência de valores negativos).

Gráfico 1 - Exemplo da distribuição dos valores cTTO



Em relação à tarefa DCE, esta componente foi monitorizada com vista a detectar padrões pouco usuais de resposta. A título de ilustração, a Tabela 2 mostra um exemplo de quantificação de seqüências de resposta tipo AAAAAAA, BBBBBBB e ABABABA.

Tabela 2 - Exemplo de quantificação dos padrões de resposta da DCE

Entrevistador	Número de Entrevistas	AAAAAAA	BBBBBBB	ABABABA	BABABAB
A	54	0	0	2	1
B	57	0	0	1	0
C	29	0	0	1	0
D	26	0	0	0	0
E	23	0	0	1	0
F	57	0	0	0	0
G	35	0	0	1	0
H	28	0	0	0	1

A periodicidade da avaliação conduzida pelo grupo coordenador nacional variou em função dos dados recolhidos, i.e., do número de entrevistas efectivamente realizadas por cada entrevistador, assumindo periodicidade diária até semanal. A avaliação realizada pela equipa de suporte do EQ-VT teve um cariz quinzenal e complementou a avaliação já efectuada pelos investigadores locais.

5.5.2. Reinquirição

Para além do processo implementado pelo grupo EuroQol, o grupo coordenador nacional estabeleceu um processo de reinquirição telefónica a um subconjunto da amostra de indivíduos.

Os objectivos deste processo de reinquirição consistiram essencialmente em garantir a ausência de fraude na realização das entrevistas, i.e., garantir que a entrevista foi efectivamente realizada e a consistência das respostas dadas. Adicionalmente, a informação desta reinquirição permitiu efectuar o teste-reteste para verificação da reprodutibilidade das respostas. No entanto, estes resultados não são analisados neste trabalho.

Foi definido pelo grupo coordenador nacional de que o processo de reinquirição seria conduzido pela empresa de sondagens a 10% dos 700 participantes previstos

entrevistar pelos entrevistadores recrutados sob a sua responsabilidade. Para efeitos de reinquirição posterior, os entrevistadores deveriam recolher informação sobre o contacto telefónico do participante e a data da realização da entrevista. As questões alvo definidas para o processo de reinquirição telefónica foram a data da realização da entrevista, o género e a idade do participante, a avaliação do estado de saúde actual de acordo com as dimensões do EQ-5D-5L e o valor atribuído na escala numerada de 0 a 100 (VAS) ao estado de saúde actual.

Por questões éticas, de forma a manter o anonimato dos participantes, estabeleceu-se que a empresa de sondagens apenas disponibilizaria um relatório final com o resultado das perguntas reinquiridas à equipa de investigadores, não facilitando qualquer acesso à informação pessoal dos respondentes.

5.6. Análise estatística

A fim de proceder à caracterização da amostra foram utilizadas estatísticas descritivas com recurso ao programa informático IBM SPSS versão 20.0. Os dados foram tratados recorrendo-se a métodos inerentes à estatística descritiva e inferencial. No que se refere à primeira, os dados foram analisados e apresentados em tabelas mostrando, nos casos aplicáveis, cálculos referentes a frequências absolutas (n .º) e relativas (%), média, desvio-padrão, mediana, percentis e valores máximos e mínimos. No estudo de hipóteses, recorreu-se ao teste de independência do Qui-Quadrado (X^2) e ao teste One Way ANOVA (F). Nos testes de hipóteses o nível de significância considerado foi de 0,10.

Para os dados de valoração obtidos através das técnicas cTTO e DCE foi também conduzida uma análise genérica baseada em estatísticas descritivas.

5.6.1. Determinação dos blocos de estados de saúde

Tendo por base os 3.125 estados de saúde gerados pelo sistema descritivo do EQ-5D-5L, o desenho para o cTTO foi iniciado com a escolha aleatória de 80 estados de saúde, a partir de um desenho factorial fraccionário obtido com recurso a um algoritmo de troca

de Fedorov, conforme indicado em Oppe et al. (2014). A estes 80 estados de saúde aleatórios acrescentaram-se seis estados específicos: o pior estado de saúde avaliado por este sistema descritivo (55555) e os estados de saúde ligeiros em que todas as dimensões assumem o nível 1 de problemas excepto uma delas que assume o segundo nível de problemas (21111, 12111, 11211, 11121 e 11112). Estes estados de saúde com alterações ligeiras foram incluídos de forma a avaliar directamente o efeito-tecto no EQ-5D-5L. Tal como definido pelos mesmos autores, considerando resultados prévios de valoração de cada um destes 86 estados de saúde, foi, de seguida, estimado um modelo de regressão a fim de calcular valores esperados para cada um destes estados de saúde. Esta estratégia foi continuada até se identificar um desenho com diferenças mínimas entre os valores prévios e os valores esperados dos estados de saúde. Por fim, os 80 estados de saúde foram distribuídos por 10 blocos de forma a que a gama completa de valores de utilidade estivesse representada em cada bloco e que cada um deles tivesse a mesma utilidade média. No desenho final, cada bloco integrou oito estados de saúde exclusivos para esse bloco, complementados com o pior estado de saúde 55555 e um dos estados ligeiros acima identificados. Daqui resultaram as 10 questões de valoração associadas no processo de entrevista à tarefa cTTO. No que concerne ao número de valorações cTTO de cada estado de saúde e considerando um número estimado de 1.000 respondentes, estimou-se obter cerca de 100 valorações para os 80 estados de saúde, 200 observações para os estados de saúde ligeiros 21111, 12111, 11211, 11121 e 11112 e 1.000 observações para o pior estado de saúde (55555).

Na componente da valoração por DCE, o desenho incluiu 196 pares de estados de saúde gerados pelo desenho experimental criado por Oppe et al. (2014). Tendo por base um desenho bayesiano eficiente, estes pares foram distribuídos por 28 blocos, cada um com sete pares de escolha (Stolk et al. 2010; Bliemer et al. 2008). Os blocos foram equilibrados em termos de nível de problemas em cada dimensão, tendo por base o índice de intensidade de problemas. Tome-se, como exemplo, o estado de saúde 22312. Neste caso, o índice de intensidade de problemas associado ao estado equivale à soma dos níveis de intensidade, i.e., $2+2+3+1+2=10$. Nenhum dos estados de saúde incluídos em cada par de escolha foi considerado dominante sobre o outro. A distribuição dos estados de saúde

pelos blocos e a posição (direita ou esquerda) do estado de saúde nos pares de escolha para valoração na tarefa DCE foram igualmente aleatórias. Com este desenho, o número de valorações de cada par de escolha, para um número estimado de 1.000 participantes, correspondeu a cerca de 36 observações.

5.6.2. Critérios de exclusão de dados

Previamente à modelação, definiram-se possíveis critérios de exclusão relativamente aos dados obtidos através da técnica cTTO, com o intuito de avaliar a qualidade da informação. A possibilidade de exclusão incidiu sobre os dados dos participantes que valoraram sempre com o mesmo valor os 10 estados de saúde atribuídos pelo EQ-VT durante a entrevista (e.g. indivíduos que não sacrificam tempo e que são designados *non-traders*), as valorações que os respondentes consideraram como incoerentes, assinaladas no módulo de feedback, os dados dos indivíduos que apresentaram um declive positivo na regressão da sua valoração e do índice de intensidade de problemas, i.e., que, em média, atribuíram valores de utilidade superiores a estados de saúde mais degradados (indivíduos classificados informalmente no estudo como *funnies*) e os estados de saúde que reflectiram inconsistências no processo de valoração, i.e., situações em que um estado de saúde descrito como menos grave foi valorado com um valor inferior relativamente ao pior estado de saúde descrito, designado como 55555.

Não foram previstos critérios de exclusão para os dados relacionados com a técnica de valoração DCE.

Tal como referido anteriormente, a fim de descrever os estados de saúde a ser valorados pelas técnicas cTTO e DCE, recorreu-se ao sistema descritivo do EQ-5D-5L. Partindo das cinco dimensões existentes neste instrumento genérico de medição em saúde – mobilidade, cuidados pessoais, actividades habituais, dor/mal-estar e ansiedade/depressão – atribuiu-se a cada dimensão cinco níveis de intensidade de problemas, que variam desde o nível 1 (ausência de problemas) ao nível 5 (problemas extremos).

Isso permitiu identificar cada estado de saúde por um número constituído por cinco dígitos, sequencialmente (da esquerda para a direita), nas dimensões mobilidade,

cuidados pessoais, actividades habituais, dor/mal-estar e ansiedade/depressão, cuja soma constitui o índice de intensidade de problemas associado a esse estado de saúde.

5.6.3. Modelação cTTO

Dado que a valoração dos estados de saúde na tarefa cTTO obedece a uma estrutura de dados em painel foram utilizadas duas abordagens para analisá-la: modelos de efeitos fixos e modelos de efeitos aleatórios.¹²

Os dados cTTO foram modelados utilizando a abordagem de efeitos fixos com o método dos mínimos quadrados ordinários (*ordinary least squares* ou OLS) e a abordagem de efeitos variáveis com o método dos mínimos quadrados generalizados (*generalized least squares* ou GLS), o modelo Tobit censurado à esquerda a -1 e o modelo de regressão por intervalo.

De acordo com Greene (2003), a principal distinção entre a abordagem de efeitos fixos e de efeitos aleatórios reside se o termo de erro incorpora ou não outros factores correlacionados com os regressores no modelo. Na abordagem de efeitos fixos assume-se que factores não observáveis são potencialmente correlacionados com os regressores, devendo ser retirados do termo de erro para utilização do método dos mínimos quadrados ordinários e explicitados no modelo pelo termo constante da regressão. Considerando que a valoração dos estados de saúde pode ser afectada por características não observáveis inerentes ao estado de saúde do indivíduo ou ao desenho do estudo e não apenas pelo nível de intensidade de problemas associado a cada dimensão, um participante pode atribuir, por exemplo, um valor inferior a um estado de saúde relativamente ao valor médio atribuído pelo resto da amostra, podendo fazê-lo da mesma forma para os restantes estados (Wittrup-Jensen et al. 2008). Isto implica que a variância do termo de erro será parcialmente determinada por esses participantes ficando, conseqüentemente, comprometida uma das suposições para a utilização do método dos mínimos quadrados ordinários (termo de erro e regressor não são correlacionados).

Tal como descrito em Dolan (1997), os coeficientes estimados pelos modelos de regressão associados a cada nível de problemas dentro de cada dimensão, permitem,

¹² Note-se que cada indivíduo valora 10 estados de saúde na técnica cTTO.

através de uma função aditiva, associar ao estado de saúde descrito pelos cinco dígitos correspondentes ao nível de problemas um índice cardinal ou valor de acordo com a seguinte fórmula:

$$V = 1 - \beta - MO - CP - AH - DM - AD$$

[10]

em que 1 representa o valor da saúde perfeita, β respeita à constante e as restantes variáveis concernem aos coeficientes estimados para cada nível de problemas nas dimensões mobilidade, cuidados pessoais, actividades habituais, dor/mal-estar e ansiedade/depressão.

a) Método dos mínimos quadrados ordinários (OLS)

A equação básica para o modelo do método dos mínimos quadrados ordinários seguiu a seguinte estrutura:

$$Y_{ij} = \beta_0 + \beta_1 MO_{ij} + \beta_2 CP_{ij} + \beta_3 AH_{ij} + \beta_4 DM_{ij} + \beta_5 AD_{ij} + \varepsilon_{ij}$$

[11]

em que Y_{ij} (variável dependente) representa a valoração do estado de saúde j através da tarefa cTTO pelo indivíduo i , β_0 representa a constante cujo significado, neste modelo, respeita ao decremento de utilidade associado a qualquer desvio da saúde perfeita, os termos MO , CP , AH , DM e AD constituem cinco regressores *dummy* correspondendo, respectivamente, às dimensões mobilidade, cuidados pessoais, actividades habituais, dor/mal-estar e ansiedade/depressão e incorporando cada uma delas os cinco níveis de problemas das dimensões do EQ-5D-5L, e ε_{ij} define o termo de erro tradicional. A equação seguinte exemplifica a variável *dummy* para a dimensão mobilidade, com os seus cinco níveis de problemas:

$$\beta_1 MO_{ij} = \beta_{1m} MO1_{ij} + \beta_{2m} MO2_{ij} + \beta_{3m} MO3_{ij} + \beta_{4m} MO4_{ij} + \beta_{5m} MO5_{ij} \quad [12]$$

Equações semelhantes aplicar-se-ão às restantes quatro dimensões incluídas no modelo.

Considerando que Y_{ij} pode ser reescrita como

$$Y_{ij} = X_{ij}\beta + \varepsilon_{ij} \quad [13]$$

e assumindo que os estimadores $\hat{\beta}$ são, sob alguma forma, função da soma do quadrado dos resíduos

$$\sum \widehat{\varepsilon}_{ij}^2 = f(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5) \quad [14]$$

no método OLS pretende-se obter estimadores $\hat{\beta}$ com o intuito de minimizar a soma do quadrado dos resíduos (Gujarati 2003):

$$\sum \widehat{\varepsilon}_{ij}^2 = \sum (Y_{ij} - \widehat{Y}_{ij})^2 = \sum (Y_{ij} - \widehat{X}_{ij}\beta)^2 \quad [15]$$

b) Método dos mínimos quadrados generalizados (GLS)

Com base no referido anteriormente, utilizou-se uma abordagem de efeitos aleatórios na qual se assumiu que a heterogeneidade não observada não está correlacionada com os regressores, não sendo por isso explicada pelo termo de erro. Neste estudo, esta abordagem foi, em primeiro lugar, consubstanciada pelo método dos mínimos quadrados generalizados que assumiu a seguinte especificação:

$$Y_{ij} = \beta_0 + \beta_1 MO_{ij} + \beta_2 CP_{ij} + \beta_3 AH_{ij} + \beta_4 DM_{ij} + \beta_5 AD_{ij} + v_{ij} + \varepsilon_{ij} \quad [16]$$

na qual

$$v_{ij} = \mu_i + \varepsilon_{ij} \quad [17]$$

Aqui, v_{ij} representa o termo de erro, μ_i constitui um termo aleatório específico de heterogeneidade individual não observada que representa em que medida o valor da constante do i -ésimo participante difere da constante geral β_0 e ε_{ij} representa factores não observáveis que afectam a variável dependente. No método GLS é minimizada a soma ponderada do quadrado dos resíduos (Gujarati 2003), ou seja:

$$\sum w_{ij} \widehat{\varepsilon}_{ij}^2 = \sum w_{ij} (Y_{ij} - \widehat{Y}_{ij})^2 = \sum w_{ij} (Y_{ij} - \widehat{X}_{ij}\beta)^2 \quad [18]$$

c) Modelo Tobit censurado à esquerda em -1

Mantendo a abordagem de efeitos aleatórios, recorreu-se ao modelo Tobit considerando uma variável latente Y_{ij}^* subjacente aos valores cTTO observados (Y_{ij}). Este modelo combina com o facto de, na tarefa cTTO, os dados estarem censurados à esquerda, i.e., em -1. O modelo Tobit explica este tipo de dados ao estimar a variável latente Y_{ij}^* , que pode tomar valores estimados para além do limite dos valores observados. Esta característica do modelo é favorável porque os valores observados são limitados pela metodologia cTTO a -1, enquanto que as preferências latentes dos participantes podem incluir valores inferiores a -1. Considerando

$$Y_{ij}^* = X_{ij} + \varepsilon_{ij} \quad [19]$$

no modelo Tobit o valor observado Y_{ij} tem as seguintes propriedades quando o valor censurado é -1:

$$Y_{ij} = \begin{cases} Y_{ij}^* & \text{se } Y_{ij}^* > -1 \\ -1 & \text{se } Y_{ij}^* \leq -1 \end{cases}$$

[20]

A equação para Y_{ij}^* é linear e semelhante à equação:

$$Y_{ij} = \beta_0 + \beta_1 MO_{1,ij} + \beta_2 CP_{2,ij} + \beta_3 AH_{3,ij} + \beta_4 DM_{4,ij} + \beta_5 AD_{5,ij} + \varepsilon_{ij}$$

[21]

com os vectores dos coeficientes do modelo a serem estimados através da função de verosimilhança do modelo Tobit. Assumindo que o termo de erro ε_{ij} segue uma distribuição normal, utiliza-se uma função de verosimilhança para ajustar os parâmetros estimados à probabilidade de Y_{ij} ser superior ao valor censurado (Johnston & DiNardo 2001):

$$L = \prod_{Y_{ij}=-1} \left[1 - \Phi \left(\frac{X_{ij}\beta}{\sigma} \right) \right] \cdot \prod_{Y_{ij}>-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \frac{(Y_{ij}-X_{ij}\beta)^2}{\sigma^2} \right]$$

[22]

d) Modelo de regressão por intervalo

A regressão por intervalo é uma generalização do modelo Tobit. Quando se sabe que o valor para o indivíduo j está dentro do intervalo $[y_{1j}; y_{2j}]$, a probabilidade que se obtém deste indivíduo é $Pr(y_{1j} \leq Y_j \leq y_{2j})$. A sua probabilidade contém termos sob a forma $Pr(Y_j \leq y_j)$ para dados censurados à esquerda e $Pr(Y_j \geq y_j)$ para dados censurados à direita, em que y_j é o valor censurado observado e Y_j constitui uma variável aleatória que representa a variável dependente do modelo. Desta forma, a regressão por intervalo permite modelar dados onde cada observação representa dados intervalares, $[a; b]$, dados censurados à esquerda, $]-\infty; b]$, ou à direita $, [a; +\infty[$, ou pontos, $a = [a; a]$. Independentemente do tipo observações, estas são consideradas na base de dados como dados intervalares.

A probabilidade do modelo de regressão por intervalo segue a dos modelos Tobit (Wooldridge 2013). Neste caso, o modelo é representado por:

$$Y_{ij}^* = X_{ij} + \varepsilon_{ij} \tag{23}$$

Para observações $j \in C$, observa-se y_j (dados tipo ponto). As observações $j \in \mathcal{L}$ e $j \in \mathcal{R}$ são, respectivamente, censuradas à esquerda e à direita, sabendo-se que a variável não observada y_j é inferior ou igual a $y_{\mathcal{L}j}$ (censurada à esquerda) ou superior ou igual a $y_{\mathcal{R}j}$ (censurada à direita), em que $y_{\mathcal{L}j}$ e $y_{\mathcal{R}j}$ são valores censurados conhecidos. As observações $j \in I$ são intervalos, sabendo-se que a variável não observada y_j encontra-se no intervalo $[y_{1j}; y_{2j}]$.

No modelo de regressão por intervalo a equação de verosimilhança assume a seguinte forma:

$$\begin{aligned} \ln L = & -\frac{1}{2} \sum_{j \in C} w_j \left\{ \left(\frac{y_j - X\beta}{\sigma} \right)^2 + \log 2\pi\sigma^2 \right\} \\ & + \sum_{j \in \mathcal{L}} w_j \log \Phi \left(\frac{y_{\mathcal{L}j} - X\beta}{\sigma} \right) \\ & + \sum_{j \in \mathcal{R}} w_j \log \left\{ 1 - \Phi \left(\frac{y_{\mathcal{R}j} - X\beta}{\sigma} \right) \right\} \\ & + \sum_{j \in I} w_j \log \left\{ \Phi \left(\frac{y_{2j} - X\beta}{\sigma} \right) - \Phi \left(\frac{y_{1j} - X\beta}{\sigma} \right) \right\} \end{aligned} \tag{24}$$

5.6.4. Modelação DCE

Os dados resultantes da tarefa DCE foram modelados assumindo utilidade aleatória, utilizando os modelos Logit e Probit. Estes modelos binomiais de escolha discreta são considerados os mais adequados pelo facto do desenho da tarefa DCE ter consistido em duas alternativas de escolha.

a) Modelo Probit

Considerando em primeiro lugar o modelo Probit, tem-se uma variável ou índice de utilidade U_{ij} que, em função da escolha exercida pelo indivíduo j , pode tomar dois valores, 0 e 1, correspondentes à escolha entre os estados de saúde A e B. No modelo Probit, define-se uma variável latente U_{ij}^* que permitirá estimar os parâmetros do índice de utilidade (U_{ij}):

$$U_{ij}^* = V_{ij} + \varepsilon_{ij} \quad [25]$$

em que

$$V_{ij} = X_{ij}\beta \quad [26]$$

A variável latente U_{ij}^* não se consegue observar, apenas se sabe que toma valores de acordo com a seguinte regra:

$$U_{ij} = \begin{cases} 1 & \text{se } U_{ij}^* > 0 \\ 0 & \text{no caso contrário} \end{cases} \quad [27]$$

Assumindo que o termo de erro ε_{ij} segue uma distribuição normal, a função normal de distribuição cumulativa é dada por:

$$F(U_{ij}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{V_{ij}} \exp^{-\frac{z^2}{2}} dz \quad [28]$$

Para obter informação sobre o índice de utilidade U_{ij} , assume-se a inversa da função normal de distribuição cumulativa (Johnston & DiNardo 2001; Hensher et al. 2015):

$$U_{ij} = F^{-1}(U_{ij}) \quad [29]$$

b) Modelo Logit

O desenvolvimento do modelo Logit é idêntico ao do Probit com exceção de que em

$$U_{ij}^* = V_{ij} + \varepsilon_{ij} \quad [30]$$

o termo de erro ε_{ij} segue o que se denomina distribuição do valor extremo ou distribuição logística. A principal diferença entre a distribuição normal e a distribuição logística é que esta última tem mais peso nos extremos (Johnston & DiNardo 2001).

Ambos os modelos incluíram os mesmos parâmetros *dummy* utilizados no modelo cTTO, reflectindo depreciações de utilidade associados aos níveis 2 a 5 para cada uma das cinco dimensões do EQ-5D-5L. O modelo de regressão assumiu a seguinte forma, considerando que j é a alternativa de escolha no conjunto de escolhas:

$$U_{ij} = \beta_1 MO_{ij} + \beta_2 CP_{ij} + \beta_3 AH_{ij} + \beta_4 DM_{ij} + \beta_5 AD_{ij} + \varepsilon_{ij} \quad [31]$$

Considerando que

$$V_{ij} = X_{ij}\beta \quad [32]$$

a função logística é formulada de acordo com o seguinte (Gujarati 2003):

$$P_{ij} = \frac{\exp^{x'_{ij}\beta}}{1 + \exp^{x'_{ij}\beta}} \quad [33]$$

5.6.5. Modelação híbrida

Por fim, foi ainda considerado um modelo híbrido que combinou os dados contínuos da tarefa cTTO com os dados dicotómicos da tarefa DCE numa função única de máxima verosimilhança, que no Stata é realizado através do comando Hyreg, desenvolvido por Ramos-Goñi et. al (2016).

Este modelo segue a seguinte especificação:

$$Y_j = x\beta + \varepsilon \quad [34]$$

na qual Y_j constitui a variável dependente do modelo, subordinada ao tipo de observação (ponto contínuo ou resposta dicotómica), β denota um vector dos coeficientes do modelo, x representa as variáveis independentes do modelo e ε o termo de erro. De acordo com os autores, a utilização de um modelo que funde respostas provenientes de diferentes técnicas, num só estimador, requer alguma precaução na medida em que a variância das respostas contínuas pode ser diferente da variância encontrada nas respostas dicotómicas e os termos de erro associados podem assumir distribuições completamente distintas.

As tarefas de valoração inerentes ao EQ-5D-5L, com resultados contínuos no caso do cTTO e discretos no caso da DCE, têm consequências para as quais, na opinião destes autores, o modelo Hyreg tem um bom desempenho. Em primeiro lugar surge a questão das respostas associadas à tarefa de valoração cTTO poderem conduzir a uma maior variabilidade nos valores obtidos para os estados de saúde piores. Isto pode suceder por o parâmetro de reescala associado às respostas cTTO poder ser proporcionalmente associado aos atributos do EQ-5D-5L, conduzindo a heteroscedasticidade. Em segundo lugar, as respostas contínuas do cTTO iguais a -1 são assumidas como iguais a este valor não permitindo respostas inferiores a -1 (i.e., foram censuradas a -1 no modelo Tobit). Por fim, as variáveis independentes nos pares de escolha da DCE representam diferenças aditivas entre os atributos das alternativas ($X_A - X_B$), tal como sucede no modelo Logit. O que os autores concluíram com a criação deste modelo é que para a modelação de uma variável de escala utilizando respostas contínuas e dicotómicas, o modelo hyreg

proporciona um enquadramento comum para as especificações da distribuição normal e logística, separa as especificações da distribuição por tipo de resposta, permite censurar pontos e limites inferiores e superiores, e possibilita o relaxamento de assumpções homoscedásticas.

Neste modelo, a equação de verosimilhança assume a seguinte forma:

$$\begin{aligned} \ln L = & -\frac{1}{2} * \sum_{j \in C} \left\{ \ln(2\pi\sigma^2) + \left(\frac{y_j - x\beta^2}{\sigma} \right)^2 \right\} \\ & + \sum_{j \in D} \left\{ \ln \left(\frac{1}{1+e^{(-x\beta/\theta)}} \right) * y_j + \ln \left(\frac{e^{(-x\beta/\theta)}}{1+e^{(-x\beta/\theta)}} \right) * (1 - y_j) \right\} \\ & + \sum_{j \in D} \left\{ \ln \left(\Phi \left(\frac{-x\beta}{\theta} \right) \right) * (1 - y_j) + \ln \left(\Phi \left(\frac{x\beta}{\theta} \right) \right) * y_j \right\} \end{aligned}$$

[35]

Em termos de implementação, o comando Hyreg estima $\ln(\sigma)$ e $\ln(\theta)$, em vez de estimar σ e θ directamente. Desta forma, os parâmetros $\ln(\sigma)$ e $\ln(\theta)$ podem ser utilizados separadamente em modelos heteroscedásticos (i.e. hetprobit ou opção het no comando intreg no Stata).

Nesta tese as análises de regressão para os modelos econométricos foram realizadas com recurso ao programa informático Stata 14.0. Os modelos analisados foram testados para a presença de heteroscedasticidade. Com vista à selecção do modelo econométrico mais adequado, definiram-se critérios para a comparação de modelos: i) consistência interna dos coeficientes; ii) significância estatística dos coeficientes; iii) critérios de informação *Akaike Information Criterion* (AIC) e *Bayesian Information Criterion* (BIC); iv) e capacidade do modelo tratar valores censurados.

5.6.6. Análise de sensibilidade

No que concerne à análise de sensibilidade, esta foi conduzida com o objectivo de avaliar o impacto da exclusão de dados associados aos estados de saúde assinalados no módulo de feedback, na modelação de dados cTTO e na combinação de dados cTTO e DCE (dados híbridos).

A análise de sensibilidade foi efectuada para verificar a robustez dos parâmetros obtidos nos modelos econométricos, analisando a consistência de resultados ao incluir todos os dados de valoração obtidos.

6. RESULTADOS

Os subcapítulos seguintes destinam-se à apresentação dos resultados deste trabalho no que respeita à supervisão da recolha de dados (6.1), à caracterização da amostra (6.2), ao dados cTTO (6.3), aos dados DCE (6.4) e à combinação dos dados obtidos com estas duas técnicas de valoração (6.5). Nas situações aplicáveis, nomeadamente na caracterização da amostra, seguiu-se a apresentação sugerida no documento *EQ-5D-5L User Guide*, publicado pelo grupo EuroQol (Rabin et al. 2015). Por fim, incluiu-se ainda uma comparação dos sistemas de valores obtidos neste estudo com os de outros países, já publicados (6.6).

6.1. Supervisão da recolha de dados

No decorrer deste estudo foram formados 54 entrevistadores. O número de pessoas formadas para a realização das entrevistas evoluiu em função das desistências ocorridas por iniciativa dos próprios e do abandono do estudo de alguns entrevistadores, por indicação do grupo coordenador nacional, devido a problemas de qualidade nas entrevistas realizadas. A formação realizada aos entrevistadores foi efectuada em grupo ou individualmente, em função do número de indivíduos recrutados em cada momento. Cada sessão formativa teve uma duração média de 4/5 horas. Dos 54 formandos, 15 indivíduos desistiram logo após esta sessão formativa, significando que 39 potenciais entrevistadores iniciaram e concluíram o teste piloto previsto no código de conduta.

Este processo formativo conduziu à realização de 253 entrevistas piloto, perfazendo uma média de 6,5 entrevistas por entrevistador. Após a realização das entrevistas piloto, executou-se o relatório de avaliação com recurso à ferramenta QC Tool, de forma a analisar detalhadamente o desempenho dos entrevistadores com vista ao apoio para ultrapassar dificuldades específicas e/ou proceder a um retreino.

A tabela 3 apresenta os resultados da avaliação do teste piloto, considerando os quatro critérios principais de qualidade definidos no processo de controlo de qualidade da tarefa cTTO de uma entrevista. De acordo com estes critérios, uma entrevista é assinalada como sendo de reduzida qualidade e, portanto, potencialmente não válida se se verificar pelo menos um dos seguintes critérios:

- i) não introdução do estado de saúde pior que morte (*lead-time* TTO) no exemplo de treino da cadeira de rodas;
- ii) tempo inferior a três minutos para executar a componente de treino do exemplo da cadeira de rodas (treino dos estados melhor e pior que morte);
- iii) inconsistência relevante na valoração dos estados de saúde no cTTO (o estado de saúde 55555 não é valorado como o pior estado e a valoração a este atribuída é pelo menos 0,5 superior à valoração do estado de saúde que obteve o pior valor);
- iv) tempo inferior a cinco minutos para a realização das 10 tarefas de valoração cTTO.

Para efeitos de apresentação de resultados, a designação dos entrevistadores foi recodificada por ordem sequencial (de 1 a 39).

Tabela 3 - Controlo de qualidade das entrevistas piloto por entrevistador e total de entrevistas

Entrevistador	Número de entrevistas	N assinaladas	% assinaladas	Cadeira de Rodas Lead-Time	% Cadeira de Rodas Lead-Time	Inconsistência relevante	% Inconsistência relevante	Tempo Cadeira de Rodas	% Tempo Cadeira de Rodas	Tempo Tarefa cTTO	% Tempo Tarefa cTTO
Ent01	5	1	20%	0	0%	0	0%	1	20%	0	0%
Ent02	5	1	20%	0	0%	0	0%	1	20%	0	0%
Ent03	8	3	38%	2	25%	0	0%	2	25%	0	0%
Ent04	5	1	20%	0	0%	0	0%	1	20%	0	0%
Ent05	5	1	20%	1	20%	0	0%	0	0%	0	0%
Ent06	6	2	33%	0	0%	0	0%	2	33%	0	0%
Ent07	5	1	20%	1	20%	0	0%	0	0%	0	0%
Ent08	5	0	0%	0	0%	0	0%	0	0%	0	0%
Ent09	5	0	0%	0	0%	0	0%	0	0%	0	0%
Ent10	5	1	20%	0	0%	0	0%	1	20%	0	0%

Entrevistador	Número de entrevistas	N assinaladas	% assinaladas	Cadeira de Rodas Lead-Time	% Cadeira de Rodas Lead-Time	Inconsistência relevante	% Inconsistência relevante	Tempo Cadeira de Rodas	% Tempo Cadeira de Rodas	Tempo Tarefa cTTO	% Tempo Tarefa cTTO
Ent11	5	1	20%	0	0%	0	0%	1	20%	0	0%
Ent12	6	2	33%	2	33%	0	0%	1	17%	0	0%
Ent13	5	0	0%	0	0%	0	0%	0	0%	0	0%
Ent14	18	7	39%	4	22%	1	6%	6	33%	0	0%
Ent15	6	2	33%	1	17%	0	0%	0	0%	1	17%
Ent16	5	0	0%	0	0%	0	0%	0	0%	0	0%
Ent17	6	6	100%	5	83%	0	0%	4	67%	1	17%
Ent18	5	4	80%	4	80%	0	0%	3	60%	2	40%
Ent19	5	4	80%	4	80%	0	0%	1	20%	0	0%
Ent20	4	4	100%	2	50%	0	0%	4	100%	1	25%
Ent21	5	4	80%	3	60%	0	0%	4	80%	1	20%
Ent22	7	7	100%	7	100%	0	0%	7	100%	0	0%
Ent23	5	3	60%	0	0%	0	0%	3	60%	0	0%
Ent24	5	5	100%	4	80%	0	0%	3	60%	1	20%
Ent25	5	1	20%	1	20%	0	0%	0	0%	0	0%
Ent26	5	2	40%	2	40%	0	0%	0	0%	0	0%
Ent27	5	5	100%	5	100%	0	0%	1	20%	0	0%
Ent28	5	4	80%	2	40%	1	20%	3	60%	1	20%
Ent29	5	0	0%	0	0%	0	0%	0	0%	0	0%
Ent30	5	2	40%	2	40%	0	0%	2	40%	0	0%
Ent31	6	1	17%	0	0%	1	17%	0	0%	0	0%
Ent32	7	4	57%	4	57%	1	14%	2	29%	0	0%
Ent33	6	2	33%	0	0%	1	17%	0	0%	1	17%
Ent34	5	0	0%	0	0%	0	0%	0	0%	0	0%
Ent35	6	2	33%	0	0%	0	0%	2	33%	0	0%
Ent36	10	5	50%	4	40%	0	0%	4	40%	1	10%
Ent37	10	3	30%	3	30%	0	0%	0	0%	0	0%
Ent38	22	18	82%	13	59%	3	14%	12	55%	8	36%
Ent39	10	5	50%	4	40%	1	10%	3	30%	1	10%
Total	253	114	45%	80	32%	9	4%	74	29%	19	8%

Nota: Os indivíduos assinalados a negrito respeitam aos entrevistadores considerados válidos para integrarem o estudo. Os assinalados a cinzento desistiram ou foram forçados a abandonar o teste piloto.

A principal dificuldade encontrada pelos entrevistadores ocorreu na fase inicial das entrevistas piloto, no exemplo da cadeira de rodas, reflectindo-se pela não introdução do estado de saúde pior que morte nesta fase de treino. Esta falha ocorreu em 80 das 253 entrevistas piloto (32%), tendo sido potenciada por uma opção pouco intuitiva do *software*

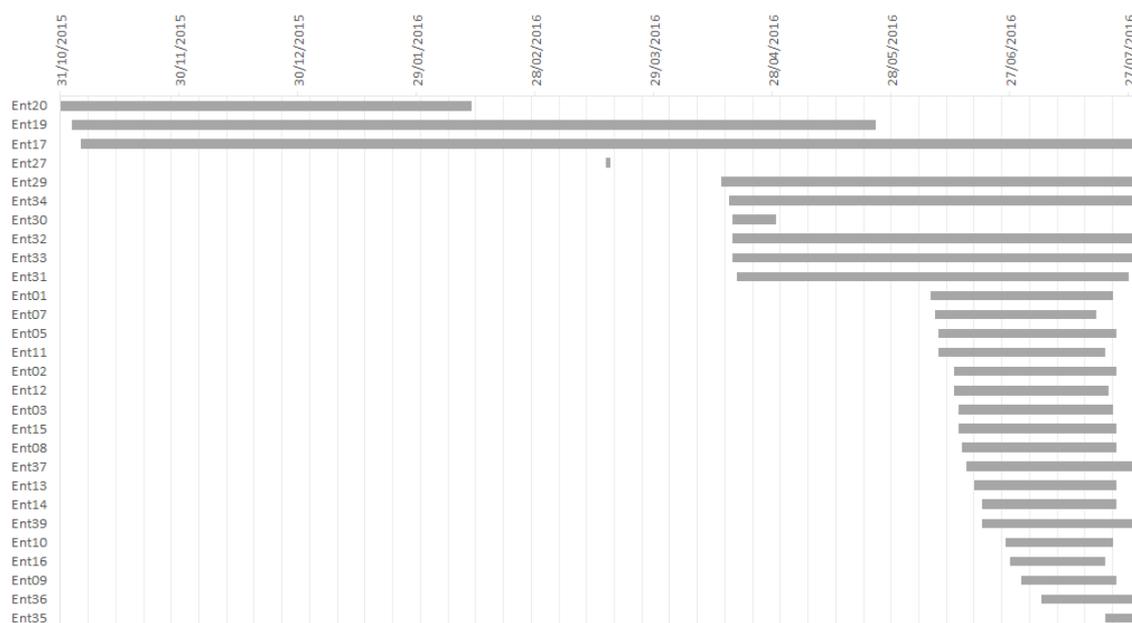
EQ-VT, que foi posteriormente resolvida com a introdução de uma correcção ao guião da entrevista e de um ênfase relativo a esta situação nas formações e acompanhamento dos entrevistadores. Pelo facto deste critério ser potenciado por circunstâncias alheias ao entrevistador, nos casos dos entrevistadores em que esta falha ocorreu maioritariamente, foi autorizado o início do trabalho de campo após retreino mas sem obrigatoriedade de repetição do teste piloto. A falha deste critério principal de qualidade teve um impacto directo noutro critério significativo, designadamente no que respeita ao tempo inferior a três minutos para a execução total do exemplo da cadeira de rodas (treino dos estados melhor e pior que morte). Esta é uma consequência óbvia na medida em que se o entrevistador falhava a introdução do estado de saúde pior que morte, dificilmente despendia o tempo mínimo estipulado na explicação de todo o exemplo da cadeira de rodas. Esta falha ocorreu em 74 entrevistas (29%). Os restantes critérios principais de qualidade não constituíram uma dificuldade acrescida para os entrevistadores. A inconsistência relevante na valoração dos estados de saúde no cTTO, i.e., quando o estado de saúde 55555 não foi valorado como o pior estado e a valoração a este atribuída foi pelo menos 0,5 superior à valoração do estado de saúde que obteve o pior valor, surgiu em 9 entrevistas (4%) e 19 entrevistas (8%) não cumpriram o tempo mínimo de cinco minutos para a concretização das 10 tarefas de valoração cTTO.

Dos 39 entrevistadores treinados e que concluíram o teste piloto, nove não chegaram a iniciar o trabalho de campo. Considerando os 30 entrevistadores restantes, dois abandonaram o estudo, por decisão da equipa coordenadora nacional por manterem um fraco desempenho na realização das entrevistas, considerando os quatro critérios principais de qualidade. Tal procedimento conduziu à perda de dados consubstanciada num total de 25 entrevistas.

Para a concretização deste estudo participaram, assim, 28 entrevistadores cujo desempenho foi considerado válido, de acordo com os critérios de qualidade estabelecidos no código de conduta. A recolha de dados efectuada pelos entrevistadores decorreu, em todo o país, durante nove meses, de 31 de Outubro de 2015 até 31 de Julho de 2016. A recolha de dados foi iniciada por três entrevistadores e finalizada por um conjunto de 24

entrevistadores activos. O período de actividade de cada um desses entrevistadores encontra-se representado no gráfico 2.

Gráfico 2 - Período de recolha de dados associado a cada entrevistador activo



Não obstante a recolha de dados ter sido extensível praticamente a todo o país, a maior parte da actividade dos entrevistadores centrou-se nos centros de recrutamento de Lisboa, Évora e Coimbra, de acordo com a distribuição apresentada na tabela 4.

Tabela 4 - Distribuição dos entrevistadores por centros de recrutamento

Entrevistadores	Lisboa	Évora	Coimbra	Total
N	12	2	14	28
%	43%	7%	50%	100%

A organização e coordenação das equipas de entrevistadores cujo trabalho de campo se desenvolveu essencialmente na zona de Évora e Lisboa foi da responsabilidade da empresa de sondagens, especialista em recolha de dados, contratada para o efeito. A organização da equipa de entrevistadores recrutada em Coimbra teve responsabilidade directa do grupo nacional de investigadores.

Durante o período de recolha de dados, realizaram-se 1.451 entrevistas consideradas válidas para integrar o estudo.

O número médio de entrevistas por entrevistador foi de 51,8 entrevistas. O tempo médio de duração das 1.451 entrevistas realizadas neste estudo foi de 49,6 minutos. A duração média mínima ocorrida nas entrevistas foi de 29,2 minutos e a máxima de 146.

Durante todo o período de recolha de dados, os entrevistadores foram sujeitos a uma avaliação contante do grupo coordenador nacional e da equipa de suporte do EQ-VT. A avaliação efectuada pela equipa coordenadora nacional consistiu na elaboração de relatórios, com recurso à ferramenta QC Tool, explicação do respectivo conteúdo a cada entrevistador e apoio à melhoria do seu desempenho. No que concerne à periodicidade da avaliação efectuada por parte dos coordenadores nacionais, esta ocorreu semanalmente nos primeiros meses de recolha de dados, passando a uma avaliação diária nos últimos dois meses de recolha de dados. Tal deveu-se ao número de entrevistas efectuadas que foi particularmente elevado nos dois meses que antecederam o final da recolha de dados e ao facto de haver necessidade de avaliar cada entrevistador de 10 em 10 entrevistas, a fim de analisar o cumprimento dos critérios principais de qualidade e de corrigir situações anómalas. De facto, nos primeiros sete meses de recolha de dados, o número médio diário de entrevistas foi de apenas 1,9 entrevistas, correspondendo a uma média de 13,2 entrevistas por semana. A actividade aumentou significativamente nos últimos dois meses de recolha de dados passando para uma média de 16,7 entrevistas diárias e 117,2 entrevistas por semana. Isso sucedeu devido ao recrutamento da equipa de entrevistadores em Coimbra que ocorreu no início de Junho de 2016.

À semelhança do que sucedeu no teste piloto, sempre que algum entrevistador teve quatro ou mais entrevistas assinaladas como não cumpridoras dos critérios principais de qualidade, em 10 entrevistas realizadas (40% ou superior), o seu trabalho de campo foi suspenso e promoveu-se uma reunião individual com os coordenadores nacionais com vista a perceber o motivo do não cumprimento e apoiar a melhoria do seu desempenho. Independentemente do resultado do retreino, as entrevistas consideradas comprometidas foram removidas da base de dados. Essa situação sucedeu com dois entrevistadores, conduzindo à remoção de um total de 18 entrevistas. A avaliação constante do desempenho permitiu a detecção precoce de falhas dos critérios principais de qualidade impedindo, assim, que os entrevistadores prosseguissem o trabalho de campo e atingissem

um número máximo de entrevistas assinaladas com qualidade comprometida. Desta forma, promoveu-se, através de retreinos preventivos, a correcção atempada do desempenho e evitou-se a perda de dados. Neste âmbito foram realizados ao longo do estudo sete retreinos de carácter individual ou em grupo. Não obstante a frequência semanal a diária da avaliação do desempenho, durante os nove meses de recolha de dados os entrevistadores receberam semanalmente dois relatórios de avaliação. Ambos os documentos continham uma explicação detalhada do desempenho individual e do grupo mas um respeitava ao número total de entrevistas realizado até ao momento (relatório total) e o outro referia-se apenas ao número de entrevistas realizado nessa semana de trabalho (relatório incremental). Desta forma, possibilitou-se uma análise da evolução global de cada entrevistador, considerando as entrevistas efectuadas desde o início do trabalho de campo, e uma avaliação mais específica, considerando apenas o período de trabalho mais recente, o que permitia a detecção de eventuais dificuldades e possíveis desvios ao guião da entrevista. Este procedimento foi essencial para garantir a qualidade das entrevistas efectuadas.

Ao longo deste período e paralelamente à avaliação efectuada pelos coordenadores nacionais, a equipa de suporte do EQ-VT elaborou relatórios quinzenais de avaliação para apoio à equipa nacional e efectuou reuniões de acompanhamento com a equipa nacional de investigadores, através de sessões remotas, em etapas pré-definidas da recolha de dados: 200, 500 e 750 entrevistas realizadas. Adicionalmente, em meados do período de recolha de dados (Abril de 2016), um dos elementos da equipa de suporte do EQ-VT deslocou-se a Portugal com o intuito de realizar uma formação adicional aos entrevistadores com o trabalho de campo em curso e a novos entrevistadores entretanto recrutados e que aguardavam ainda a formação inicial. Esta sessão intensiva teve a duração de dois dias e incidiu sobre o conteúdo da entrevista, o processo de controlo de qualidade e simulações de entrevistas. O grupo coordenador nacional participou activamente na organização desta reunião e no apoio às sessões práticas.

Durante os nove meses de recolha de dados, os coordenadores nacionais providenciaram apoio diário aos entrevistadores, relativamente a dificuldades do foro informático durante a utilização da plataforma EQ-VT. Nos casos em que a equipa nacional

não conseguiu solucionar os problemas colocados pelos entrevistadores, promoveram-se sessões remotas com um dos elementos da equipa de suporte do EQ-VT em Roterdão, Holanda, e os entrevistadores afectados.

Tal como previamente definido, a empresa de sondagens reinquiriu um total de 148 participantes, correspondentes a 10,2% dos indivíduos da amostra.

6.2. Caracterização da amostra

6.2.1. Representatividade

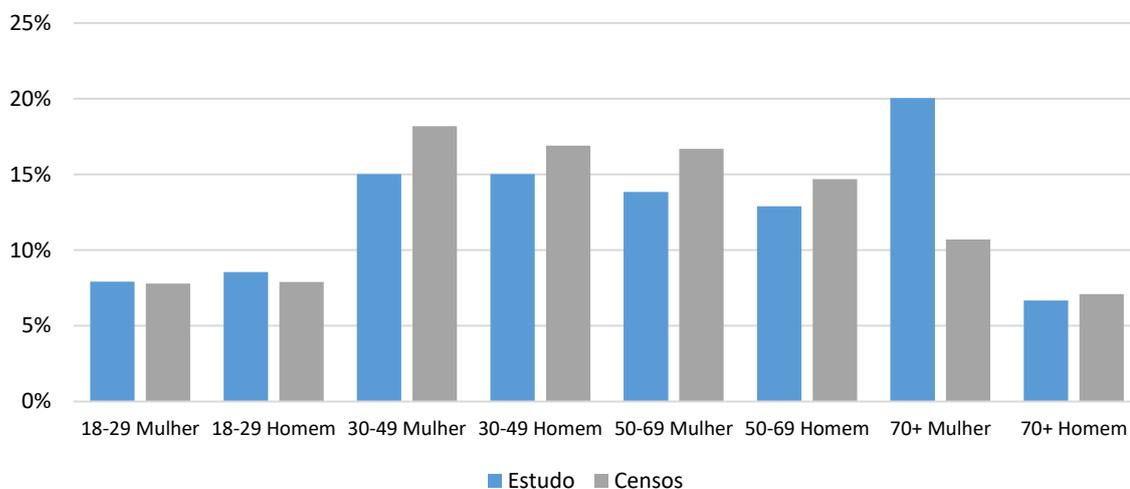
Tal como referido anteriormente, a amostra foi projectada de forma a obter-se um conjunto de indivíduos representativos da população portuguesa no que concerne ao género e idade. Os cálculos subjacentes à definição da amostra tiveram por base a distribuição da população portuguesa descrita no Censos de 2011 para um conjunto de 1.000 indivíduos.

Considerando que tiveram a sua entrevista validada neste estudo 1.451 indivíduos, a tabela 5 e o gráfico 3 comparam a distribuição por género e grupos etários da amostra projectada, a partir dos dados do Censos de 2011, com a amostra participante no estudo.

Tabela 5 - Distribuição da amostra projectada pelo Censos de 2011 e da amostra real do estudo, por género e grupo etário

Género	Masculino		Feminino		Total	
	Censos	Estudo	Censos	Estudo	Censos	Estudo
Grupo Etário N(%)						
18-29 anos	79(7,9%)	124(8,5%)	78(7,8%)	115(7,9%)	157(15,7%)	239(16,5%)
30-49 anos	169(16,9%)	218(15,0%)	182(18,2%)	218(15,0%)	352(35,2%)	436(30,0%)
50-69 anos	147(14,7%)	187(12,9%)	167(16,7%)	201(13,9%)	314(31,4%)	388(26,7%)
>=70 anos	71(7,1%)	97(6,7%)	107(10,7%)	291(20,1%)	178(17,8%)	388(26,7%)
Total	467(46,7%)	626(43,1%)	533(53,3%)	825(56,9%)	1.000(100,0%)	1.451(100,0%)

Gráfico 3 - Comparação dos grupos por grupo etário e género: amostra projectada pelo Censos de 2011 versus estudo



Comparando a amostra projectada pelo Censos de 2011 com a amostra real do estudo, verifica-se que não pode ser rejeitada a hipótese nula da amostra real ser representativa da população portuguesa de referência ($\chi^2(0,05; 3) = 0,0000$; $p < 0,001$).

6.2.2. Análise descritiva da amostra

A amostra foi constituída por 1.451 indivíduos residentes em Portugal.

A distribuição da amostra segundo o género mostra uma maioria de indivíduos do sexo feminino (56,9%) em relação ao sexo masculino (Tabela 6). Quanto à idade, e considerando os grupos etários definidos no documento *EQ-5D-5L User Guide*, observa-se que o grupo com maior número de indivíduos corresponde ao que tinha idade igual ou superior a 70 anos (26,7%), seguido das faixas etárias dos 18 aos 29 anos (16,5%), dos 30 aos 39 anos (16,3%) e dos 50 aos 59 anos (16,1%). O grupo etário com menor número de indivíduos foi o dos 60 aos 69 anos (10,7%). A média de idades foi de $51,2 \pm 19,4$ anos, sendo a idade mínima de 18 anos e a máxima de 92 anos.

Tabela 6 - Características sociodemográficas da amostra

Característica Sociodemográfica	N	%
Género		
Feminino	825	56,9
Masculino	626	43,1
Grupo Etário		
[18; 29] anos	239	16,5
[30; 39] anos	236	16,3
[40; 49] anos	200	13,8
[50; 59] anos	233	16,1
[60; 69] anos	155	10,7
≥70 anos	388	26,7
Nível de Escolaridade		
Não sabe ler e escrever	10	0,7
Sabe ler e escrever sem possuir grau de ensino	37	2,5
1.º ciclo do ensino básico	239	16,5
2.º ciclo do ensino básico	86	5,9
3.º ciclo do ensino básico	176	12,1
Ensino secundário	322	22,2
Ensino pós-secundário	80	5,5
Ensino politécnico ou universitário	501	34,1
Estado Civil		
Solteiro(a)	409	28,2
Casado(a) ou em união de facto	732	50,4
Divorciado(a) ou separado(a)	137	9,4
Viúvo(a)	173	11,9
Situação Laboral		
Empregado(a)/Auto-emprego	726	50,0
Desempregado(a) mas capaz de trabalhar	70	4,8
Desempregado(a) mas incapaz de trabalhar por motivos de saúde	3	0,2
Estudante	146	10,1
Doméstico(a)	43	3,0
Reformado(a)/Aposentado(a)	463	31,9
Rendimento Médio Mensal Líquido do Agregado		
<500€	116	11,4
[500;1.000[€	224	22,0
[1.000;2.000[€	303	29,7
[2.000;3.000[€	182	17,8
≥3.000€	107	10,5
Dependente do agregado familiar	88	8,6

Não obstante cerca de um terço dos indivíduos deter um grau de ensino politécnico ou universitário (34,1%), a amostra incluiu ainda uma minoria de indivíduos que não sabe ler e escrever (0,7%). No entanto, considerando a amostra em análise, pode observar-se que a maioria dos indivíduos (61,8%) possuía um grau de ensino igual ou superior ao secundário.

Quanto ao estado civil da amostra, observa-se que a maioria dos indivíduos era casada ou vivia em união de facto (50,4%), 28,2% dos indivíduos referia ser solteiro, 11,9% era viúvo e apenas 9,4% era divorciado ou separado.

No que respeita à situação laboral, observa-se que metade dos indivíduos entrevistados (50%) encontrava-se empregado, assumindo-se como trabalhador dependente ou independente. 31,9% dos indivíduos referiu ser reformado ou aposentado e apenas 10,1% eram estudantes, não exercendo qualquer actividade laboral.

O rendimento médio mensal líquido do agregado familiar que integra o participante encontrava-se entre [1.000; 2.000[€ para 29,7% dos indivíduos. Segue-se 22% dos indivíduos cujo agregado familiar auferia rendimentos mensais contidos no intervalo [500; 1.000[€. No entanto, salienta-se que quase um terço da amostra não respondeu a esta questão, por desconhecer ou por intencionalmente não pretender disponibilizar essa informação.

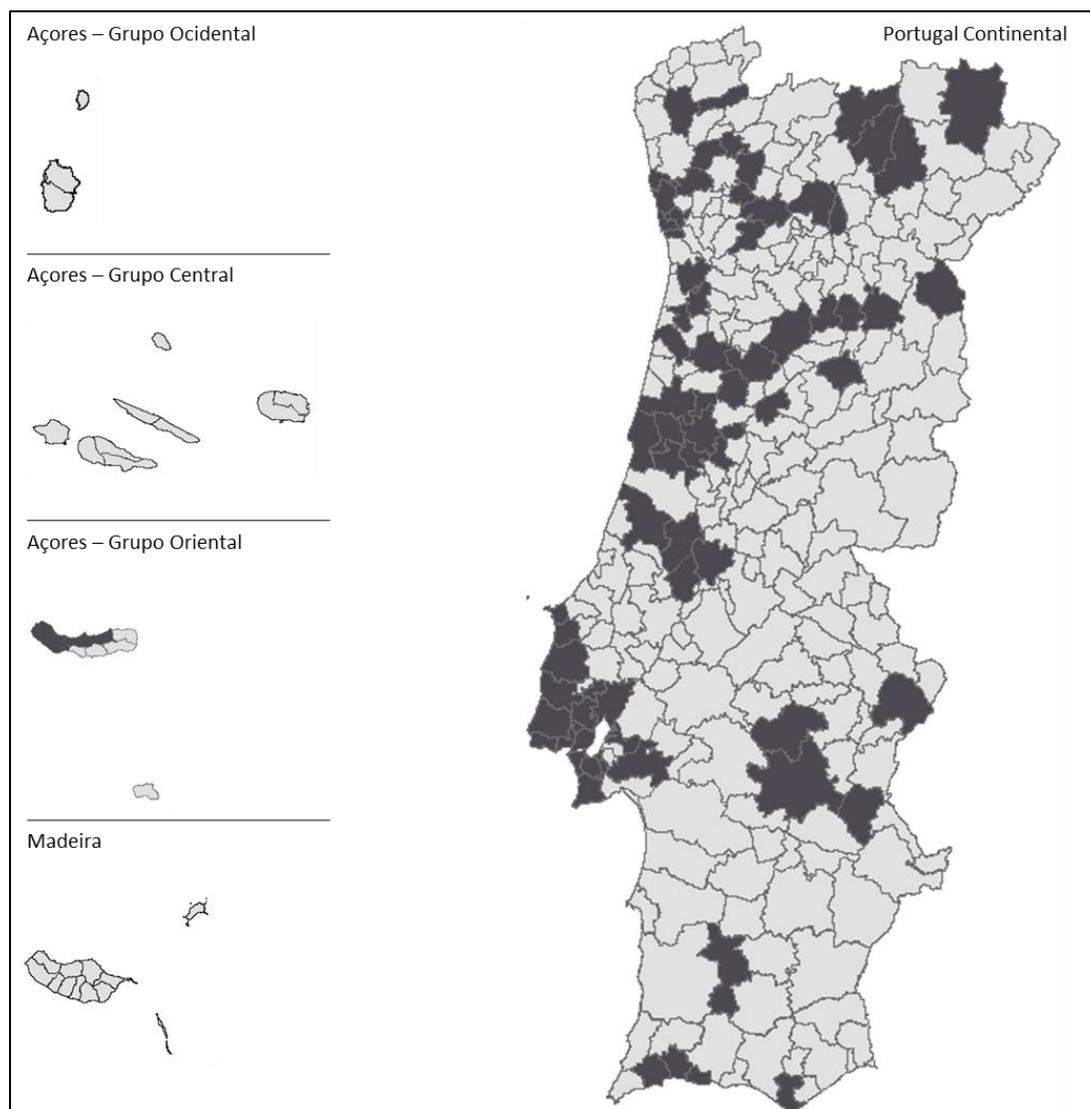
Relativamente à dimensão do agregado familiar, o número médio de elementos correspondeu a 3 ± 1 . O número mínimo foi de 1 e o máximo de 8.

No que diz respeito ao local de residência, verifica-se que a maior parte dos indivíduos entrevistados residia em freguesias classificadas como predominantemente urbanas (1.228; 85,4%). Apenas 9,3% (134) residia em freguesias predominantemente rurais e 5,3% (76) em áreas maioritariamente urbanas.

Fazendo a associação da freguesia ao concelho respectivo, verifica-se que os participantes residiam em 79 concelhos distintos pertencentes a Portugal continental e ao grupo oriental de ilhas que integram o arquipélago dos Açores.

De forma a possibilitar uma melhor visualização, apresenta-se a figura 42 que assinala a negrito os concelhos de residência dos participantes. Os concelhos com maior número de participantes foram Coimbra, com 402 entrevistados (27,7%), Évora, com 228 participantes (15,7%), e Lisboa, com 206 indivíduos (14,2%).

Figura 42 - Distribuição da amostra por concelho de residência do entrevistado



Passando agora para a análise do estado de saúde da amostra (Tabela 7), a maior parte dos indivíduos (62,3%) referiu não ter qualquer doença que tenha sido comunicada por um profissional de saúde. Para os indivíduos que responderam afirmativamente a esta pergunta (32,9%), perguntou-se se podiam identificar a doença que foi diagnosticada. As patologias mais referidas, de forma isolada ou associadas a outras doenças, foram a hipertensão arterial (n=47), diabetes (n=38) e patologia cardíaca (n=38), doença oncológica (n=25), alergias (n=16) e asma (n=11).

Tabela 7 - Estado de saúde/Experiência com doença do entrevistado

Estado de saúde/Experiência com doença	Amostra (N)	Amostra (%)
Doença comunicada por profissional de saúde		
Sim	478	32,9
Não	904	62,3
Não sabe/Não responde	69	4,8
Auto-avaliação da saúde		
Muito má	2	0,1
Má	29	2,0
Razoável	494	34,0
Boa	727	50,1
Muito boa	199	13,7
Ocorrência de doença grave		
Sim	329	22,7
Não	1.122	77,3
Experiência com familiar com doença grave		
Sim	1.063	77,3
Não	388	26,7
Experiência como cuidador de pessoa com doença crónica		
Sim	496	34,2
Não	955	65,8

No decorrer da entrevista foi pedido ao indivíduo que fizesse uma auto-avaliação qualitativa do seu estado de saúde. Neste âmbito, metade dos entrevistados (50,1%) referiu ter um estado de saúde que considera ser bom, 34% referiu ter uma saúde razoável e 13,7% respondeu que considera a sua saúde muito boa. Apenas 2,1% dos indivíduos considerou ter um estado de saúde mau ou muito mau.

Outra das perguntas efectuadas visava a identificação de contacto prévio com doenças graves, do entrevistado ou de algum familiar, e de experiência prévia do entrevistado como cuidador de uma pessoa com doença crónica. De acordo com a informação apresentada no quadro 11, a maioria dos entrevistados (77,3%) não teve o que considera ser uma doença grave. No entanto, 77,3% dos indivíduos já teve contacto com uma situação de doença grave de um familiar e 34,2% já cuidou de uma pessoa com doença crónica. Daqui conclui-se que, não obstante o contacto com uma doença grave não ser a título individual, a maioria dos entrevistados já experienciou de perto essa situação.

De seguida analisa-se o estado de saúde actual auto-reportado dos indivíduos que integram a amostra, utilizando o sistema descritivo do EQ-5D-5L e a VAS.

A tabela 8 mostra a frequência relativa do nível de problemas em cada dimensão do sistema descritivo do EQ-5D-5L, de acordo com a auto-avaliação dos indivíduos da amostra relativamente ao seu estado de saúde no momento da entrevista.

Tabela 8 - Frequência relativa dos problemas por dimensão e grupo etário

Dimensão	Nível	18-29	30-39	40-49	50-59	60-69	≥70	Total
Mobilidade	Sem problemas	97,9	94,9	92,5	85,8	70,3	32,7	74,4
	Problemas ligeiros	2,1	3,8	6,5	8,2	20,0	34,3	14,5
	Problemas moderados	0,0	1,3	0,5	4,3	7,1	26,3	8,8
	Problemas graves	0,0	0,0	0,5	1,3	2,6	6,4	2,3
	Problemas extremos	0,0	0,0	0,0	0,4	0,0	0,3	0,1
Cuidados Pessoais	Sem problemas	98,7	99,6	99,0	94,8	89,7	61,9	87,5
	Problemas ligeiros	1,3	0,4	1,0	1,7	6,5	28,1	8,9
	Problemas moderados	0,0	0,0	0,0	3,0	2,6	8,0	2,9
	Problemas graves	0,0	0,0	0,0	0,4	1,3	1,5	0,6
	Problemas extremos	0,0	0,0	0,0	0,0	0,0	0,5	0,1
Actividades Habituais	Sem problemas	94,6	93,2	94,0	84,1	72,9	39,2	75,5
	Problemas ligeiros	4,6	6,8	4,5	9,4	20,0	29,9	14,1
	Problemas moderados	0,8	0,0	1,5	4,3	4,5	25,8	8,4
	Problemas graves	0,0	0,0	0,0	2,1	2,6	3,9	1,7
	Problemas extremos	0,0	0,0	0,0	0,0	0,0	1,3	0,3
Dor/ Mal-estar	Sem problemas	76,2	76,3	69,0	59,7	43,9	18,8	53,8
	Problemas ligeiros	20,5	22,5	27,5	28,8	36,8	42,5	30,7
	Problemas moderados	2,5	1,3	3,5	6,9	14,2	33,0	12,5
	Problemas graves	0,4	0,0	0,0	4,3	5,2	5,7	2,8
	Problemas extremos	0,4	0,0	0,0	0,4	0,0	0,0	0,1
Ansiedade/ Depressão	Sem problemas	72,0	79,7	73,5	70,8	63,2	53,6	67,4
	Problemas ligeiros	20,9	17,4	21,0	19,3	25,8	33,0	23,8
	Problemas moderados	5,9	2,5	5,0	6,9	9,0	11,1	7,1
	Problemas graves	1,3	0,4	0,5	1,7	1,9	1,8	1,3
	Problemas extremos	0,0	0,0	0,0	1,3	0,0	0,5	0,3

Analisando de uma forma global o quadro anterior, conclui-se que as dimensões que apresentaram um maior número de indivíduos que referiu não ter problemas foram as dimensões funcionais: “cuidados pessoais” (87,5%), “actividades habituais” (75,5%) e “mobilidade” (74,4%). Apesar de superior a 50% mas com um menor número de indivíduos, surgiram as dimensões “ansiedade/depressão” (67,4%) e “dor/mal-estar” (53,8%). É também interessante observar que a percentagem de indivíduos sem problemas decresceu com a idade. As dimensões com maior ocorrência de problemas extremos (0,3%) respeitaram às “actividades habituais” e à “ansiedade/depressão”.

De forma a facilitar a análise por dimensão e segundo as recomendações do grupo EuroQol, englobaram-se os indivíduos que referiram apresentar algum tipo de problema (nível 2 a nível 5). Neste âmbito, a tabela 9 apresenta o número e percentagem de indivíduos com e sem problemas, por dimensão e grupo etário.

Tabela 9 - Frequência absoluta e relativa de indivíduos com e sem problemas, por dimensão e grupo etário

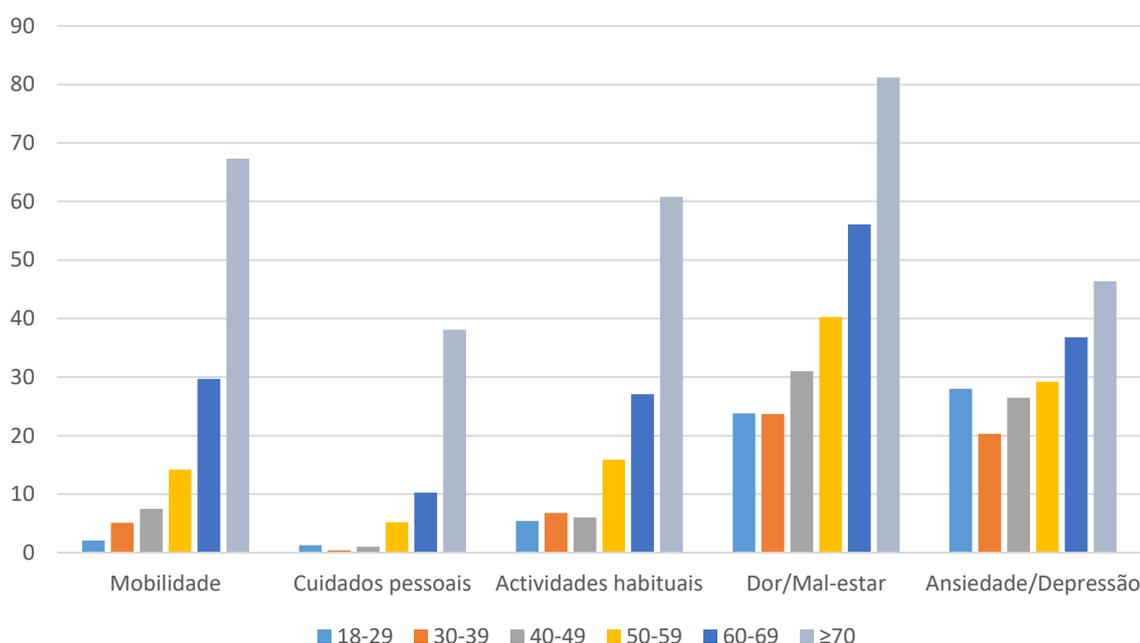
Dimensão	Problemas		18-29	30-39	40-49	50-59	60-69	≥70	Total	χ^2
Mobilidade	Sem problemas	N	234	224	185	200	109	127	1.079	526,4 (p<0,001)
		%	97,9	94,9	92,5	85,8	70,3	32,7	74,4	
	Com problemas	N	5	12	15	33	46	261	372	
		%	2,1	5,1	7,5	14,2	29,7	67,3	25,6	
Cuidados pessoais	Sem problemas	N	236	235	198	221	139	240	1.269	327,8 (p<0,001)
		%	98,7	99,6	99,0	94,8	89,7	61,9	87,5	
	Com problemas	N	3	1	2	12	16	148	182	
		%	1,3	0,4	1,0	5,2	10,3	38,1	12,5	
Actividades habituais	Sem problemas	N	226	220	188	196	113	152	1.095	410,3 (p<0,001)
		%	94,6	93,2	94,0	84,1	72,9	39,2	75,5	
	Com problemas	N	13	16	12	37	42	236	356	
		%	5,4	6,8	6,0	15,9	27,1	60,8	24,5	
Dor/ Mal-estar	Sem problemas	N	182	180	138	139	68	73	780	315,0 (p<0,001)
		%	76,2	76,3	69,0	59,7	43,9	18,8	53,8	
	Com problemas	N	57	56	62	94	87	315	671	
		%	23,8	23,7	31,0	40,3	56,1	81,2	46,2	
Ansiedade/ Depressão	Sem problemas	N	172	188	147	165	98	208	978	57,9 (p<0,001)
		%	72,0	79,7	73,5	70,8	63,2	53,6	67,4	
	Com problemas	N	67	48	53	68	57	180	473	
		%	28,0	20,3	26,5	29,2	36,8	46,4	32,6	

Na análise da relação entre cada dimensão e os grupos etários, os valores do teste χ^2 e de p permitem concluir que há diferenças estatisticamente significativas entre estas variáveis. Essa diferença é explicada por um número superior ao esperado de indivíduos com idade igual ou superior a 70 anos com maior número de problemas reportados dentro de cada dimensão.

O gráfico 4 apresenta a soma das percentagens correspondentes aos níveis 2 (problemas ligeiros) a 5 (problemas extremos) para cada uma das cinco dimensões do EQ-5D-5L para os seis grupos etários em análise. Como pode observar-se, a percentagem de

problemas reportados aumentou com a progressão da idade. Neste âmbito, pode observar-se que os indivíduos mais idosos (idade igual ou superior a 70 anos) reportaram mais problemas em todas as dimensões. O efeito da idade é particularmente notório na dimensão mobilidade, em que se observa um aumento de 37,6 pontos percentuais do grupo etário dos 60 aos 69 anos para o de idade igual ou superior a 70 anos. O menor aumento ocorreu na dimensão ansiedade/depressão com 9,6 pontos percentuais.

Gráfico 4 - Percentagem de problemas reportados por grupo etário



A análise de dados ao estado de saúde actual da amostra evidencia ainda que 41,4% dos indivíduos referiu não ter qualquer problema em todas as dimensões do sistema descritivo do EQ-5D-5L (estado de saúde classificado como 11111) e 58,6% indicou problemas em pelo menos uma das dimensões (Tabela 10).

Tabela 10 - Estado de saúde global auto-reportado por grupo etário

Estado de Saúde		18-29	30-39	40-49	50-59	60-69	≥70	Total
11111	N	145	145	117	110	42	41	600
	%	60,7	61,4	58,5	47,2	27,1	10,6	41,4
Todos os outros	N	94	91	83	123	113	347	851
	%	39,3	38,6	41,5	52,8	72,9	89,4	58,6

Na análise aos resultados das 1.451 avaliações da VAS observa-se que o valor médio obtido foi de 78,9 (numa escala de 0 a 100) com um desvio-padrão de 14,9. O valor mínimo foi de 1 e o máximo foi de 100. A tabela 11 apresenta as estatísticas descritivas da VAS, por grupo etário e para o total da amostra. Tal como pode também observar-se, os valores médios da VAS diminuíram à medida que a idade progrediu. A análise ao desvio-padrão não permite verificar uma evolução consistente com a idade mas é notório que este valor aumenta para os grupos etários com idade igual ou superior a 40 anos, significando uma maior variabilidade nos resultados da VAS.

Tabela 11 – Estatísticas descritivas da VAS por grupo etário

VAS	18-29	30-39	40-49	50-59	60-69	≥70	Total
Média	87,0	87,2	84,7	79,1	76,3	66,8	78,9
Desvio-Padrão	9,5	9,3	10,4	15,5	14,3	13,9	14,9
Mediana	90,0	90,0	90,0	90,0	80,0	70,0	80,0
Percentil 25	80,0	80,0	80,0	70,0	70,0	55,0	70,0
Percentil 75	95,0	92,8	90,0	90,0	86,0	76,8	90,0

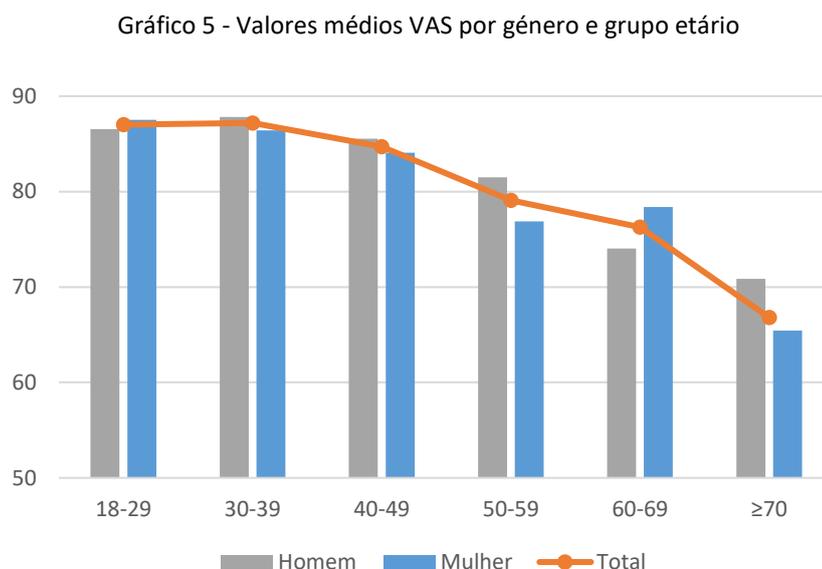
Quando se analisa a amostra por grupos, considerando a pontuação obtida relativamente ao nível de problemas identificados na avaliação do estado de saúde actual com recurso ao sistema descritivo do EQ-5D-5L, observa-se que, à medida que o índice de intensidade de problemas aumentou, o valor médio atribuído na VAS diminuiu e aumentou o desvio-padrão (Tabela 12). De notar que na avaliação auto-reportada de saúde não se encontraram avaliações cuja pontuação se situa entre 20 e 25 e que equivalem à descrição de estados de saúde caracterizados por um elevado nível de problemas.

Tabela 12 - Estatísticas da VAS por pontuação obtida nos níveis de problemas do estado de saúde auto-reportado

Índice de intensidade de problemas:	Média	Desvio-Padrão
5 (ausência de problemas)	87,3	8,7
[6; 10]	76,6	13,4
[11; 15]	58,0	13,0
[16; 20]	53,6	17,3
[21; 25]	--	--

Quando se compara a relação entre os grupos que resultam das distintas pontuações atribuídas ao índice de intensidade de problemas e a VAS, conclui-se que há diferenças estatisticamente significativas entre estas duas variáveis ($F = 295,16; p < 0,001$).

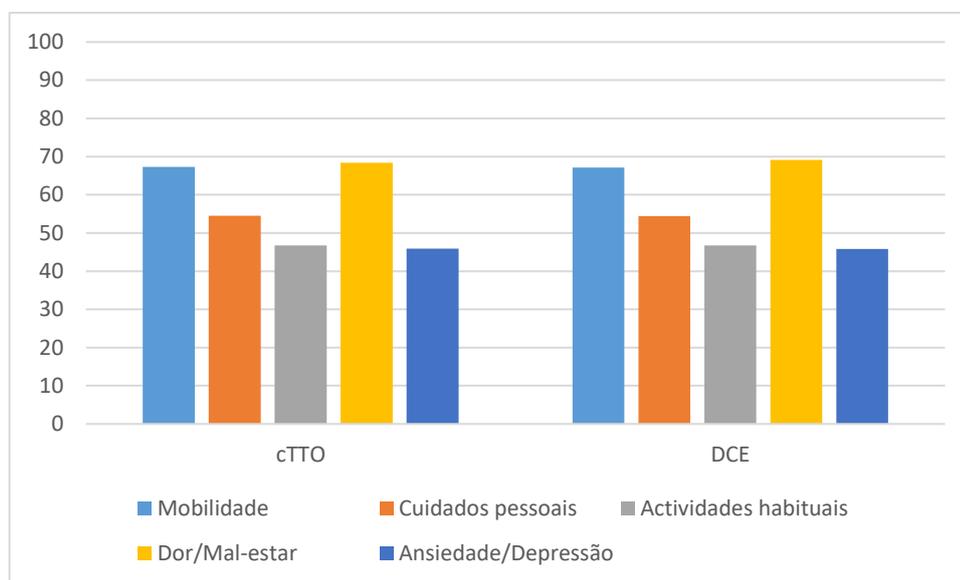
Para finalizar a análise dos resultados da VAS, o gráfico 5 apresenta os valores médios da VAS por género e grupo etário, permitindo observar o que já anteriormente havia sido referido relativamente aos valores médios da VAS decrescerem com a progressão da idade mas acrescentando que, nos grupos etários entre os 30 e os 59 anos de idade e nas idades iguais ou superiores a 70 anos, os homens apresentaram um valor médio da VAS superior ao das mulheres. Tal já não sucedeu entre os 18 e os 29 anos e os 60 e os 69 anos.



Apresenta-se de seguida os resultados relativamente às dimensões do sistema descritivo do EQ-5D-5L que os indivíduos da amostra consideraram mais relevantes durante a execução das tarefas de valoração cTTO e DCE. Tal como pode observar-se no gráfico 6, independentemente da técnica de valoração utilizada (cTTO ou DCE), os indivíduos atribuíram maior relevância à dimensão dor/mal-estar (68,4% e 69,1%, respectivamente) seguida da mobilidade (67,3% e 67,1%) e dos cuidados pessoais (54,5%

e 54,4%). A dimensão menos relevante foi, em ambos os casos, a ansiedade/depressão (45,9% e 45,8%).

Gráfico 6 - Relevância das dimensões do sistema descritivo do EQ-5D-5L



Por fim, identifica-se na tabela 13 o número de participantes que, no processo de valoração cTTO, não trocou tempo, i.e., não participou no processo de equivalência de tempo de vida ou qualidade de vida, valorando todos os estados de saúde iguais a 1, sendo designados, como já referido, por *non-traders*. Adicionalmente, identifica-se o número de participantes que valoraram todos os estados de saúde de forma igual (mas distinta de 1) e os indivíduos cujos dados apresentaram um declive positivo na regressão entre a sua valoração e o índice de intensidade dos problemas, i.e., que, em média, atribuíram valores de utilidade superiores relativamente a estados de saúde mais degradados (indivíduos classificados informalmente no estudo como *funnies*).

Tabela 13 - Número e percentagem de indivíduos que valorou todos os estados de saúde de forma igual e de *funnies*

Participantes:	N	%
Valorações iguais a 1 (<i>non-traders</i>)	5	0,3%
Valorações iguais a 0,95	3	0,2%
Valorações iguais a -1	2	0,2%
<i>Funnies</i>	1	0,1%

6.3. Modelo cTTO

6.3.1. Controlo de qualidade

No que concerne ao desempenho dos 28 entrevistadores com entrevistas válidas que integraram este estudo, apresentam-se os resultados obtidos relativamente ao incumprimento, individual ou simultâneo, dos quatro critérios principais de qualidade na tarefa cTTO, que poderão indiciar entrevistas de qualidade reduzida (Tabela 14).

Tabela 14 - Controlo de qualidade das entrevistas válidas (n=1.451)

Número entrevistas assinaladas	% assinaladas	Cadeira de rodas <i>lead-time</i>	% Cadeira de rodas <i>lead-time</i>	Inconsistência relevante	% Inconsistência relevante	Tempo cadeira de rodas	% Tempo cadeira de rodas	Tempo tarefa cTTO	% Tempo tarefa cTTO
135	9%	34	2%	12	1%	90	6%	18	1%

De forma a facilitar a leitura do quadro, relembram-se os quatro critérios principais de qualidade definidos no processo de controlo de qualidade da tarefa cTTO de uma entrevista que a suceder, individualmente ou em grupo, obrigam a assinalar uma entrevista como sendo de reduzida qualidade e, portanto, potencialmente não válida: i) não introdução do estado de saúde pior que morte (*lead-time* TTO) no exemplo de treino da cadeira de rodas; ii) tempo inferior a três minutos para executar a componente de treino do exemplo da cadeira de rodas (treino dos estados melhor e pior que morte); iii) inconsistência relevante na valoração dos estados de saúde no cTTO (o estado de saúde 55555 não é valorado como o pior estado e a valoração a este atribuída é pelo menos 0,5 superior à valoração do estado de saúde que obteve o pior valor); e iv) tempo inferior a cinco minutos para a realização das 10 tarefas de valoração cTTO.

Tal como pode observar-se, de um total de 1.451 entrevistas válidas, 135 (9%) foram assinaladas devido ao incumprimento de um ou mais critérios principais. A maior dificuldade residiu no critério relacionado com o tempo despendido no treino do exemplo da cadeira de rodas (treino dos estados de saúde melhor e pior que morte) que definia um tempo mínimo de três minutos na execução total da tarefa para evitar que a entrevista

fosse assinalada. Neste caso específico, 90 entrevistas (6%) foram assinaladas como tendo um tempo inferior a três minutos associado ao treino do exemplo da cadeira de rodas. Com o segundo maior número de ocorrências surge a não introdução do estado de saúde pior que morte (*lead-time*) no exemplo da cadeira de rodas. Nestas circunstâncias estão 34 entrevistas, correspondentes a 2% do total de entrevistas válidas realizadas. O incumprimento deste critério influencia directamente o incumprimento do tempo total despendido no exemplo da cadeira de rodas. Esta situação torna-se óbvia na medida em que se uma das partes do treino é suprimida (introdução do estado de saúde pior que morte) o tempo total para desempenhar a tarefa será menor. Por fim, surgem as inconsistências flagrantes na valoração do estado de saúde 55555 que ocorreram em 12 (1%) entrevistas e a duração inferior a cinco minutos na realização das 10 tarefas de valoração cTTO que surgiram em 18 (1%) entrevistas.

De uma forma geral, os entrevistadores cumpriram os critérios principais de qualidade alcançando menos de 40% de entrevistas assinaladas no total de entrevistas realizadas.

Outro dos aspectos analisados pela ferramenta QC Tool correspondeu ao desempenho dos entrevistadores no exemplo de treino da cadeira de rodas. A tabela 15 apresenta o tempo médio da execução da totalidade do treino da cadeira de rodas, por entrevistador, bem como o tempo médio associado às componentes de treino referentes ao estado de saúde melhor que morte e ao estado de saúde pior que morte.

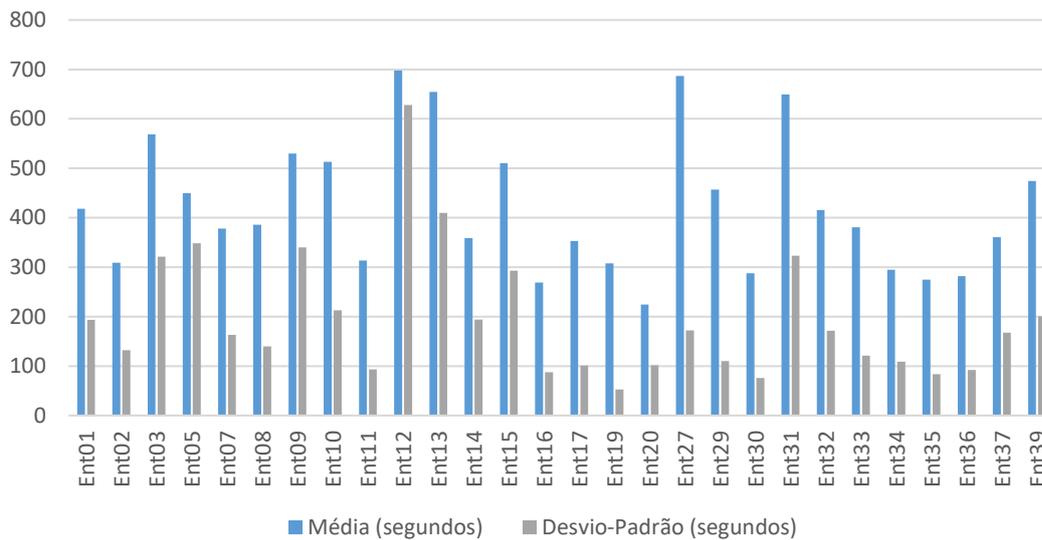
Tabela 15 - Duração média do exemplo da cadeira de rodas (n=1.451)

Tempo médio exemplo cadeira de rodas (segundos)	Desvio-padrão exemplo cadeira de rodas (segundos)	Tempo médio estado melhor que morte no exemplo cadeira de rodas (segundos)	Desvio-padrão estado melhor que morte no exemplo cadeira de rodas (segundos)	Tempo médio estado pior que morte no exemplo cadeira de rodas (segundos)	Desvio-padrão estado pior que morte no exemplo cadeira de rodas (segundos)
421,60	135,94	281,66	101,22	139,94	58,51

Os resultados evidenciam que a duração média do tempo dedicado ao exemplo da cadeira de rodas excedeu, em todos os entrevistadores, os três minutos (180 segundos) mínimos preconizados pelo protocolo do estudo (Gráfico 7). Considerando o desempenho

da totalidade dos entrevistadores, a duração média foi de 421,60 ± 135,94 segundos. A duração média mínima foi de 224,19 segundos e a máxima foi de 697,69 segundos.

Gráfico 7 - Duração média do exemplo da cadeira de rodas, por entrevistador



O protocolo prevê que cerca de dois terços do tempo total do exemplo da cadeira de rodas seja despendido na explicação do estado de saúde melhor que morte e que um terço seja dedicado ao estado de saúde pior que morte. Não obstante esta última parte representar uma dificuldade acrescida para o participante, dado tratar-se de um conceito potencialmente mais complexo de entender, ainda assim, é previsível que mais tempo seja despendido na primeira parte dada a necessidade de se mostrar a forma de funcionamento da plataforma EQ-VT no exemplo da cadeira de rodas. Os dados mostram que, em média, 281,66 segundos (dois terços do tempo total da tarefa) foram destinados à explicação do estado de saúde melhor que morte e que 139,94 segundos (um terço do tempo total da tarefa) foram reservados ao estado de saúde pior que morte. Os gráficos 8 e 9 apresentam o tempo médio dedicado por cada entrevistador, respectivamente, ao estado de saúde melhor que morte e pior que morte no exemplo da cadeira de rodas.

Gráfico 8 - Duração média da explicação do estado de saúde melhor que morte no exemplo da cadeira de rodas, por entrevistador

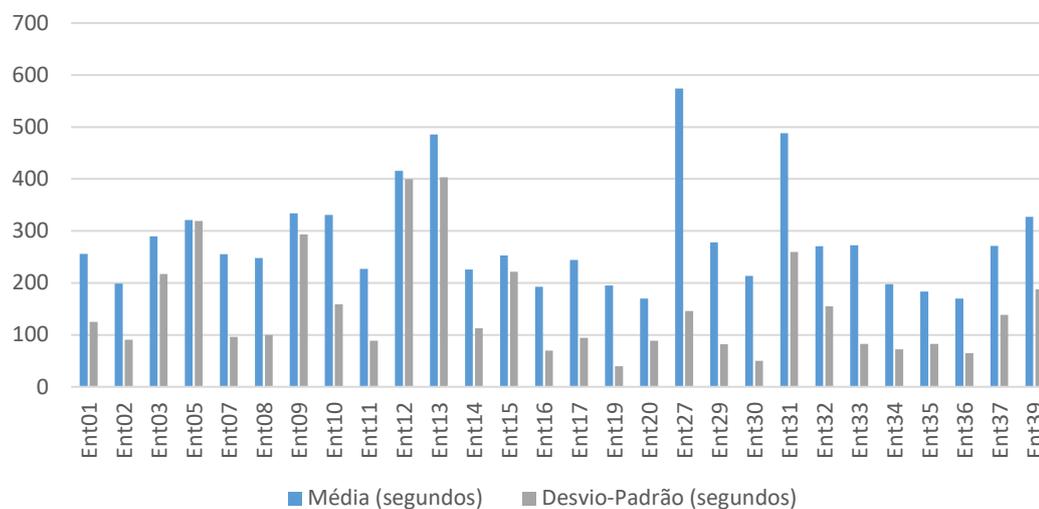
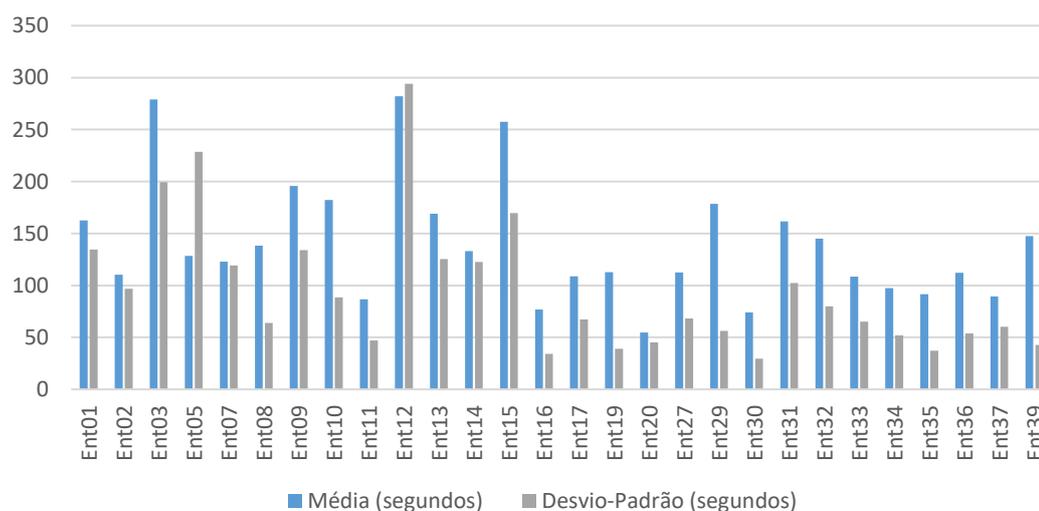


Gráfico 9 - Duração média da explicação do estado de saúde pior que morte no exemplo da cadeira de rodas, por entrevistador



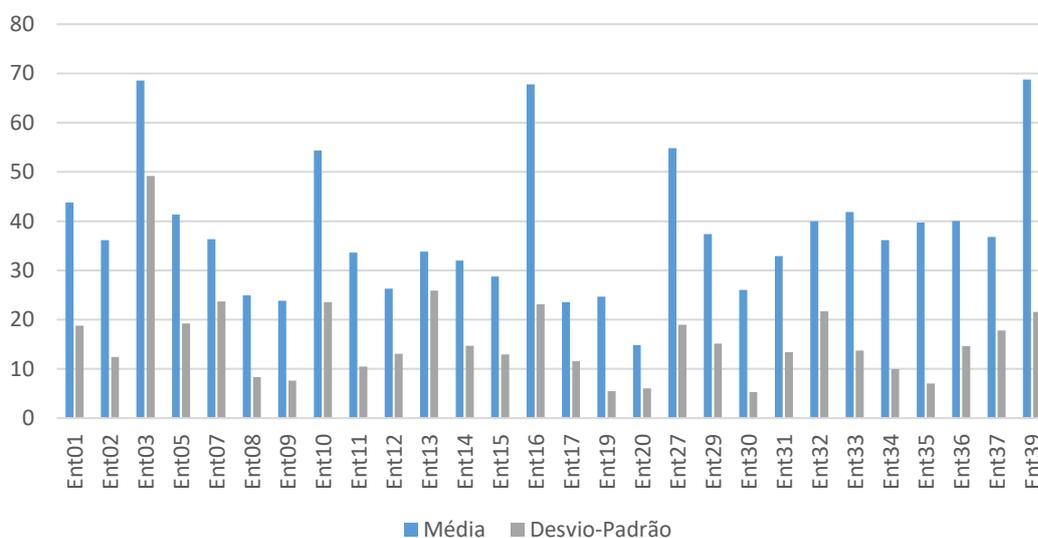
Ainda durante esta fase de treino, a ferramenta QC Tool monitorizou o número de movimentos efectuados pelo entrevistador no processo iterativo de valoração do exemplo da cadeira de rodas (Tabela 16). A lógica subjacente a esta monitorização reside no facto de se entender que, quanto maior for o número de movimentos, melhor será a apresentação ao participante do funcionamento da técnica de valoração e da panóplia possível de respostas e, assim, obter uma resposta mais válida.

Tabela 16 - Número médio de movimentos no exemplo da cadeira de rodas (n=1.451)

Movimentos médios exemplo cadeira de rodas	Desvio-padrão exemplo cadeira de rodas	Movimentos médios estado melhor que morte no exemplo cadeira de rodas	Desvio-padrão estado melhor que morte no exemplo cadeira de rodas	Movimentos médios estado pior que morte no exemplo cadeira de rodas	Desvio-padrão estado pior que morte no exemplo cadeira de rodas
38	14	25	10	13	6

Não obstante não existir um número mínimo de movimentos associado ao processo iterativo de valoração no exemplo da cadeira de rodas, o protocolo considera que, de forma a explicar consistentemente o método, deverão ser efectuados pelo menos 30 movimentos. Neste âmbito, apenas 28% dos entrevistadores não atingiu este objectivo. O gráfico 10 apresenta a informação relativa ao número médio de movimentos por entrevistador.

Gráfico 10 - Número médio de movimentos no exemplo da cadeira de rodas, por entrevistador



Neste caso e ao invés do que se preconiza relativamente à distribuição do tempo na explicação dos estados de saúde melhor e pior que morte, o número de movimentos associado ao estado de saúde pior que morte deverá ser superior ao número de movimentos consignado ao estado de saúde melhor que morte.

Na verdade, este objectivo não foi alcançado pelos entrevistadores, possivelmente pela dificuldade em superar o elevado número de movimentos já realizados no estado melhor que morte para introduzir a tarefa.

Os gráficos 11 e 12 apresentam a informação relativa ao número de movimentos iterativos executados no exemplo da cadeira de rodas nas componentes associadas ao estado de saúde melhor e pior que morte, respectivamente.

Gráfico 11 – Número médio de movimentos no estado de saúde melhor que morte no exemplo da cadeira de rodas, por entrevistador

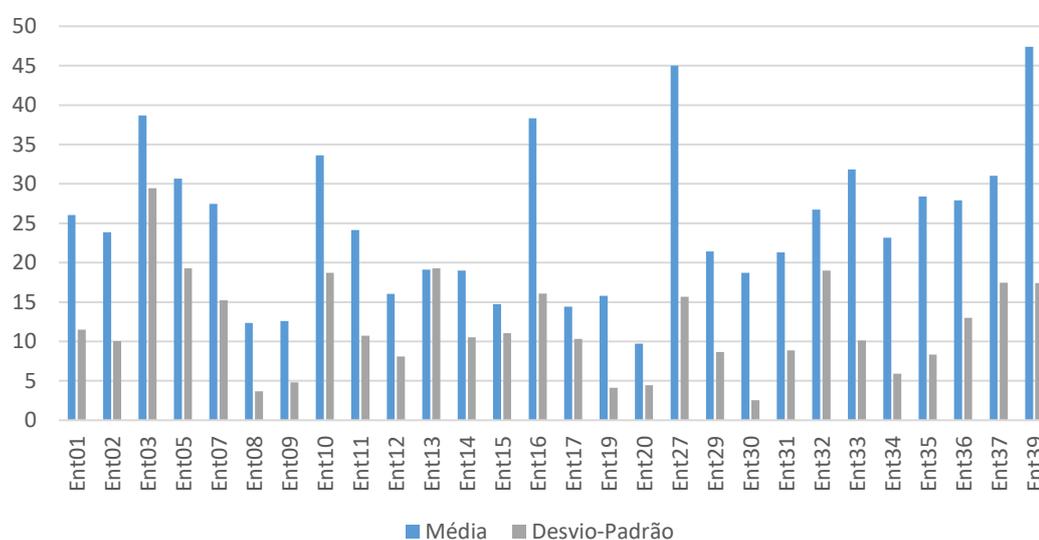
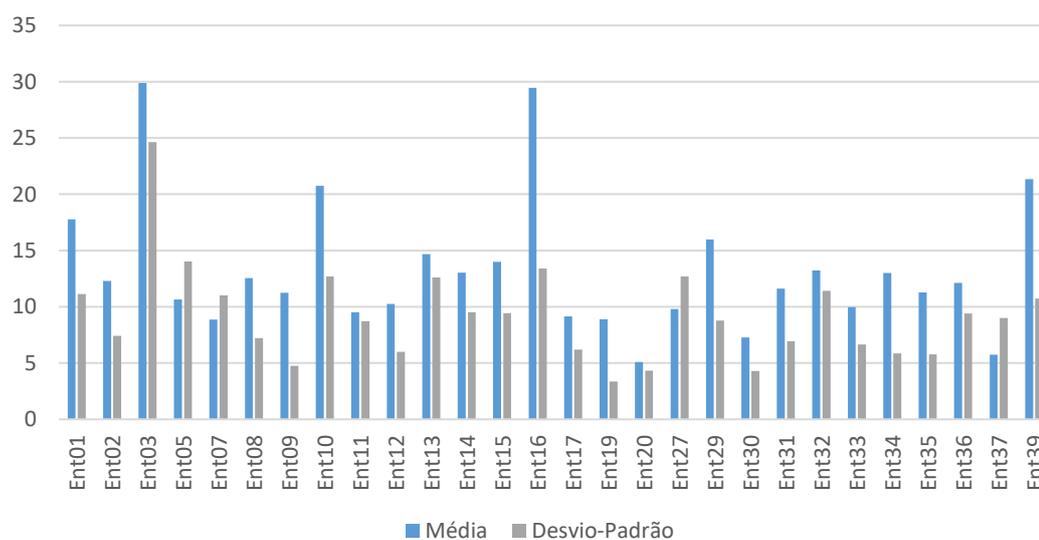


Gráfico 12 - Número médio de movimentos no estado de saúde pior que morte no exemplo da cadeira de rodas, por entrevistador



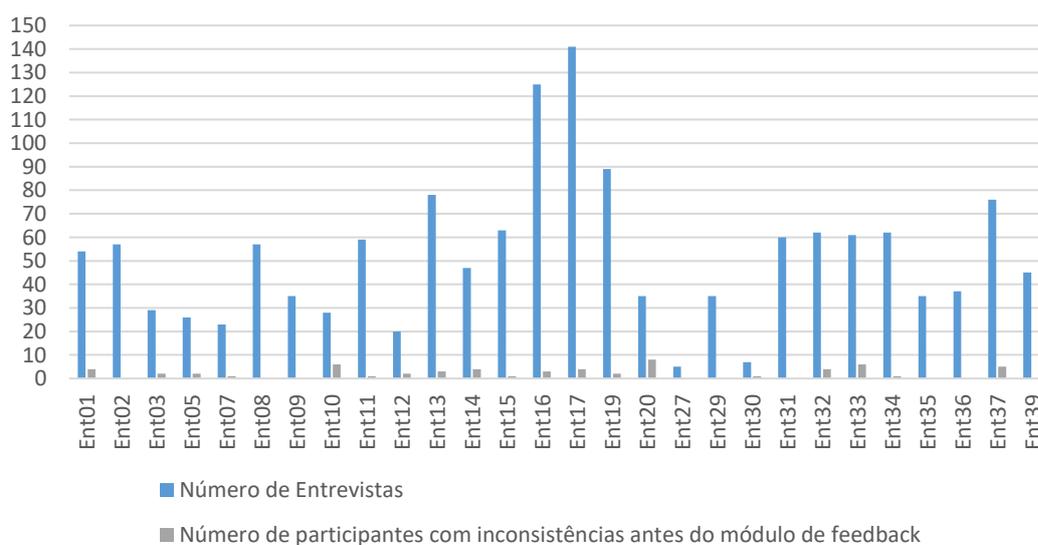
Outro dos aspectos avaliados na adesão ao protocolo por parte dos entrevistadores foi a promoção da utilização do módulo de feedback por parte dos participantes (Tabela 17). Neste âmbito, apenas um entrevistador não utilizou o módulo de feedback como forma de promover, junto do participante, a revisão da ordenação das respostas.

Tabela 17 - Estados de saúde assinalados no módulo de feedback (n=1.451)

Número de estados valorados	Número de estados assinalados	Percentagem de estados assinalados	Tempo médio módulo de feedback (segundos)	Desvio-padrão módulo de feedback (segundos)
14.510	1.119	8%	232,70	194,38

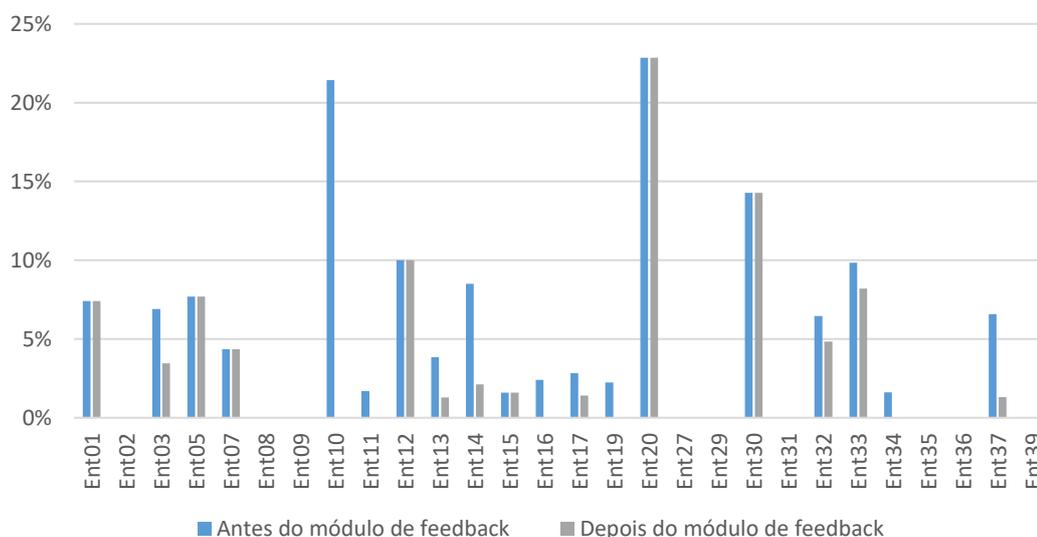
A principal vantagem da utilização do módulo de feedback decorre da possibilidade do participante identificar estados de saúde cuja ordenação não considera correcta. Desta forma, os estados assinalados pelo participante poderão ser excluídos pelos investigadores para efeitos de modelação dos dados. A análise ao Gráfico 13 permite observar o número de participantes cujos dados de valoração cTTO continham pelo menos uma inconsistência, relacionada como estado de saúde 55555, face ao número total de entrevistas realizadas por entrevistador.

Gráfico 13 - Número de participantes com inconsistências relativamente ao número total de entrevistas



O gráfico 14 permite observar a evolução da percentagem de participantes com inconsistências antes e depois da utilização do módulo de feedback, onde tiveram a oportunidade de assinalar estados de saúde considerados com ordenação incorrecta. Lembra-se que os 10 estados de saúde valorados por cada participante eram apresentados no módulo de feedback por ordem decrescente de valoração.

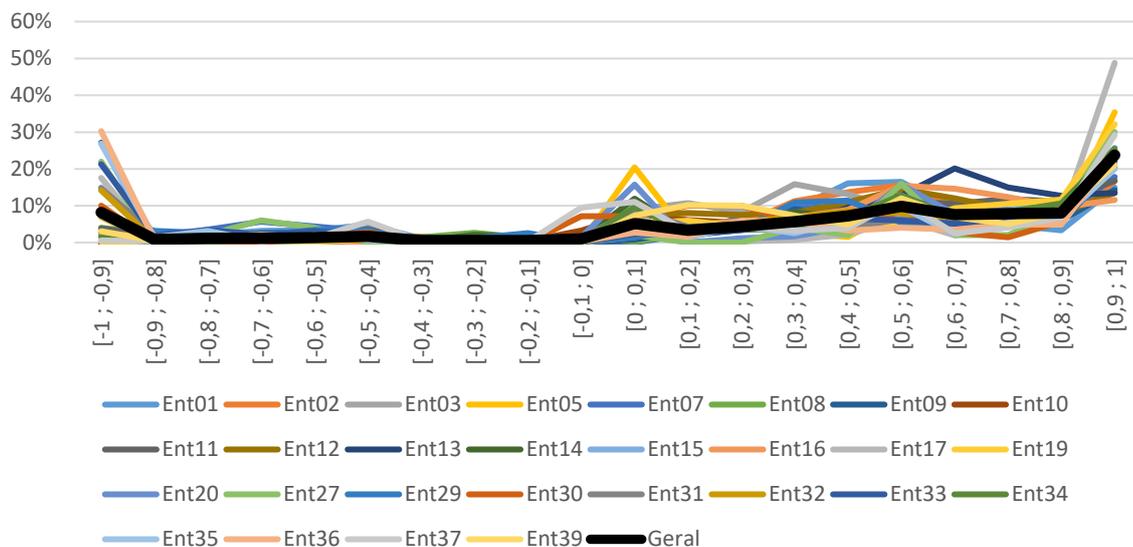
Gráfico 14 - Percentagem de participantes com inconsistências antes e depois da utilização do módulo de feedback



Prévia à utilização do módulo de feedback identificaram-se 60 participantes com inconsistências envolvendo o estado de saúde 55555. Após a sua utilização, o número de participantes com inconsistências reduziu-se para 33.

Com o objectivo de avaliar a existência de poucas valorações piores que morte (negativas) associadas a um entrevistador e *clusters* (picos) em torno de certos valores tais como 1; 0,5; 0; -0,5 e -1 apresenta-se, no gráfico 15, a distribuição de valores cTTO associada a cada entrevistador. Tal como pode observar-se, a distribuição de valores dos entrevistadores acompanhou a distribuição dos valores gerais, não se verificando a existência de picos marcados.

Gráfico 15 - Distribuição de valores cTTO por entrevistador



6.3.2 Análise dos dados

Neste estudo, cada um dos 1.451 participantes contribuiu com 10 valorações reais de estados de saúde com recurso à técnica cTTO, resultando num total de 14.510 valorações. Cada uma das valorações cTTO foi executada, em média, em 78,6 segundos (1,3 minutos). A tabela 18 apresenta as estatísticas descritivas para os valores cTTO obtidos nos 86 estados de saúde em análise.

Tabela 18 - Estatísticas descritivas dos valores cTTO

Estado de Saúde	N	Média	Desvio-Padrão	Estado de Saúde	N	Média	Desvio-Padrão	Estado de Saúde	N	Média	Desvio-Padrão
11112	306	0,95	0,13	21345	133	0,38	0,47	43315	146	0,39	0,49
11121	287	0,95	0,09	21444	152	0,36	0,48	43514	133	0,31	0,50
11122	143	0,91	0,13	22434	147	0,50	0,42	43542	145	0,19	0,52
11211	292	0,94	0,08	23152	133	0,43	0,44	43555	143	-0,13	0,57
11212	133	0,91	0,19	23242	160	0,59	0,43	44125	135	0,29	0,56
11221	139	0,94	0,07	23514	145	0,41	0,48	44345	135	-0,03	0,58
11235	139	0,59	0,46	24342	160	0,49	0,46	44553	133	-0,05	0,53
11414	152	0,61	0,38	24443	146	0,36	0,47	45133	145	0,28	0,55
11421	151	0,82	0,17	24445	143	0,09	0,56	45144	139	0,15	0,56
11425	147	0,49	0,50	24553	147	0,17	0,50	45233	151	0,38	0,49

Estado de Saúde	N	Média	Desvio-Padrão	Estado de Saúde	N	Média	Desvio-Padrão	Estado de Saúde	N	Média	Desvio-Padrão
12111	282	0,95	0,08	25122	151	0,66	0,37	45413	147	0,37	0,47
12112	133	0,90	0,20	25222	152	0,63	0,41	51152	147	0,29	0,47
12121	145	0,93	0,07	25331	152	0,57	0,43	51451	139	0,13	0,58
12244	151	0,45	0,45	31514	152	0,47	0,43	52215	145	0,25	0,59
12334	160	0,60	0,43	31524	146	0,40	0,50	52335	143	0,20	0,55
12344	135	0,34	0,51	31525	151	0,40	0,45	52431	146	0,39	0,52
12513	135	0,62	0,37	32314	160	0,60	0,48	52455	151	0,00	0,52
12514	139	0,46	0,44	32443	145	0,34	0,51	53221	135	0,53	0,50
12543	145	0,30	0,51	33253	160	0,34	0,53	53243	152	0,27	0,54
13122	147	0,84	0,14	34155	145	0,07	0,59	53244	152	0,14	0,56
13224	143	0,57	0,45	34232	143	0,56	0,42	53412	160	0,43	0,55
13313	151	0,77	0,22	34244	133	0,30	0,49	54153	146	0,08	0,55
14113	146	0,75	0,28	34515	139	0,24	0,54	54231	139	0,38	0,54
14554	135	-0,06	0,57	35143	152	0,36	0,47	54342	135	0,12	0,58
15151	146	0,27	0,54	35245	139	0,15	0,58	55225	160	0,12	0,60
21111	284	0,94	0,13	35311	143	0,51	0,49	55233	151	0,22	0,54
21112	135	0,92	0,11	35332	147	0,49	0,43	55424	133	-0,01	0,51
21315	146	0,56	0,46	42115	147	0,50	0,46	55555	1451	-0,45	0,50
21334	160	0,61	0,45	42321	143	0,67	0,38	--	--	--	--

Analisando os valores cTTO numa perspectiva de intensidade do nível de problemas, i.e., da pontuação resultante da soma dos níveis de problemas nas cinco dimensões que caracterizam os estados de saúde obtêm-se os resultados apresentados na tabela 19.

Tabela 19 - Valores médios e desvio-padrão por índice de intensidade de problemas dos estados de saúde valorados pelo cTTO

Índice de intensidade de problemas	Média	Desvio-Padrão
6	0,95	0,11
7	0,92	0,14
9	0,83	0,16
10	0,75	0,28
11	0,69	0,32
12	0,61	0,42
13	0,52	0,47
14	0,45	0,46

Índice de intensidade de problemas	Média	Desvio-Padrão
15	0,39	0,51
16	0,34	0,51
17	0,33	0,49
18	0,16	0,56
19	0,10	0,57
20	-0,02	0,55
21	-0,03	0,52
22	-0,13	0,57
25	-0,45	0,50

Os dados cTTO foram modelados excluindo apenas os estados de saúde assinalados no módulo de feedback. Esta decisão foi tomada pelo grupo coordenador nacional na medida em que o módulo de feedback tem por objectivo permitir ao participante a revisão das suas respostas, identificando estados de saúde considerados mal ordenados. Desta forma, as 14.510 valorações de estados de saúde obtidas através da técnica cTTO reduziram-se para 13.391, excluindo-se, assim, 1.119 valorações de estados de saúde consideradas pelos respondentes como incorrectamente ordenados. Não foram excluídas outras situações previstas nos critérios de exclusão possíveis elencados anteriormente. No que respeita aos indivíduos que não abdicam de tempo, classificados como *non-traders* (n=5), assumiu-se que a existência desses indivíduos, para além de ser normal numa amostra representativa da população, apresentou uma ínfima expressão neste estudo. Quanto aos indivíduos que valoraram todos os estados de saúde com o mesmo valor, não obstante denotar uma compreensão reduzida da tarefa de valoração ou desinteresse no processo, o facto de se tratar apenas de 5 indivíduos motivou a que não fossem excluídos. O mesmo sucedeu com o único indivíduo que, em média, atribuiu valores de utilidade superiores relativamente a estados de saúde mais degradados, designado como *funny*. Por fim, na medida em que parte das inconsistências ocorridas em 60 participantes foi solucionada no módulo de feedback, optou-se por não se efectuar qualquer exclusão destes participantes.

Considerando os critérios previamente definidos para a selecção dos modelos econométricos, designadamente a consistência interna e a significância estatística dos coeficientes, o ajustamento do modelo tendo por base os critérios de informação AIC e BIC

e o tratamento de valores censurados, as tabelas 20 e 21 apresentam os resultados dos modelos utilizados na modelação de dados cTTO e o desempenho de cada um deles de acordo com os critérios previamente definidos.¹³

Tabela 20 - Resultados e avaliação dos modelos OLS, GLS efeitos variáveis e GLS censurado (n=13.391)

EQ-5D	Modelo 1: OLS			Modelo 2: GLS efeitos variáveis			Modelo 3: GLS censurado		
	β	Erro-Padrão	p	β	Erro-Padrão	p	β	Erro-Padrão	p
mo2	0,006	0,013	0,631	0,012	0,010	0,240	0,008	0,011	0,483
mo3	0,071	0,014	0,000	0,072	0,011	0,000	0,066	0,012	0,000
mo4	0,175	0,016	0,000	0,166	0,012	0,000	0,165	0,013	0,000
mo5	0,298	0,014	0,000	0,298	0,011	0,000	0,308	0,012	0,000
cp2	0,042	0,013	0,002	0,030	0,010	0,004	0,027	0,011	0,013
cp3	0,084	0,015	0,000	0,078	0,012	0,000	0,078	0,012	0,000
cp4	0,182	0,015	0,000	0,160	0,012	0,000	0,164	0,013	0,000
cp5	0,261	0,014	0,000	0,261	0,010	0,000	0,278	0,011	0,000
ah2	0,028	0,014	0,046	0,032	0,011	0,003	0,032	0,011	0,006
ah3	0,068	0,015	0,000	0,069	0,012	0,000	0,069	0,012	0,000
ah4	0,144	0,015	0,000	0,152	0,011	0,000	0,154	0,012	0,000
ah5	0,231	0,014	0,000	0,229	0,011	0,000	0,243	0,011	0,000
dm2	0,026	0,013	0,040	0,023	0,010	0,016	0,020	0,010	0,052
dm3	0,088	0,015	0,000	0,092	0,012	0,000	0,092	0,013	0,000
dm4	0,246	0,013	0,000	0,239	0,010	0,000	0,247	0,011	0,000
dm5	0,371	0,015	0,000	0,373	0,011	0,000	0,389	0,012	0,000
ad2	0,027	0,014	0,048	0,033	0,011	0,003	0,031	0,012	0,009
ad3	0,068	0,016	0,000	0,082	0,012	0,000	0,078	0,013	0,000
ad4	0,210	0,015	0,000	0,220	0,011	0,000	0,223	0,012	0,000
ad5	0,273	0,014	0,000	0,279	0,011	0,000	0,291	0,011	0,000
constante	0,033	0,014	0,019	0,030	0,014	0,026	0,031	0,015	0,036
Inconsistências	0			0			0		
Sem significância estatística ($\alpha=10\%$)	1			1			1		
AIC	16.946,135			12.133,876			15.491,875		
BIC	17.103,684			12.306,460			15.664,429		

Nota: mo – mobilidade; cp – cuidados pessoais; ah – actividades habituais; dm – dor/mal-estar; ad – ansiedade/depressão

¹³ Para cada um dos coeficientes estimados foram calculados também intervalos de confiança e testada a presença de heteroscedasticidade nos modelos.

Tabela 21 - Resultados e avaliação dos modelos Tobit censurado, Intervalo GLS censurado em 1 (GLS_1) e Intervalo GLS censurado em 1 e 0 (GLS_1_0) (n=13.391)

EQ-5D	Modelo 4: Tobit censurado			Modelo 5: Intervalo GLS_1			Modelo 6: Intervalo GLS_1_0		
	β	Erro-Padrão	p	β	Erro-Padrão	p	β	Erro-Padrão	p
mo2	0,002	0,015	0,909	0,008	0,011	0,483	0,005	0,011	0,637
mo3	0,065	0,015	0,000	0,066	0,012	0,000	0,060	0,012	0,000
mo4	0,175	0,017	0,000	0,165	0,013	0,000	0,164	0,013	0,000
mo5	0,308	0,015	0,000	0,308	0,012	0,000	0,321	0,012	0,000
cp2	0,041	0,014	0,004	0,027	0,011	0,013	0,025	0,011	0,025
cp3	0,086	0,016	0,000	0,078	0,012	0,000	0,077	0,013	0,000
cp4	0,188	0,016	0,000	0,164	0,013	0,000	0,166	0,013	0,000
cp5	0,277	0,015	0,000	0,278	0,011	0,000	0,293	0,012	0,000
ah2	0,028	0,015	0,065	0,032	0,011	0,006	0,030	0,012	0,011
ah3	0,070	0,016	0,000	0,069	0,012	0,000	0,069	0,013	0,000
ah4	0,145	0,016	0,000	0,154	0,012	0,000	0,156	0,012	0,000
ah5	0,244	0,015	0,000	0,243	0,011	0,000	0,257	0,012	0,000
dm2	0,023	0,014	0,084	0,020	0,010	0,052	0,019	0,011	0,072
dm3	0,089	0,017	0,000	0,092	0,013	0,000	0,091	0,013	0,000
dm4	0,254	0,014	0,000	0,247	0,011	0,000	0,256	0,011	0,000
dm5	0,386	0,016	0,000	0,389	0,012	0,000	0,403	0,012	0,000
ad2	0,026	0,015	0,082	0,031	0,012	0,009	0,029	0,012	0,015
ad3	0,062	0,017	0,000	0,078	0,013	0,000	0,075	0,014	0,000
ad4	0,213	0,016	0,000	0,223	0,012	0,000	0,227	0,013	0,000
ad5	0,284	0,015	0,000	0,291	0,011	0,000	0,302	0,012	0,000
constante	0,033	0,015	0,031	0,031	0,015	0,036	0,031	0,015	0,040
Inconsistências	0			0			0		
Sem significância estatística ($\alpha=10\%$)	1			1			1		
AIC	20.506,664			15.491,875			15.887,626		
BIC	20.671,716			15.664,429			16.060,180		

Nota: mo – mobilidade; cp – cuidados pessoais; ah – actividades habituais; dm – dor/mal-estar; ad – ansiedade/depressão

A análise às tabelas 21 e 22 permite verificar que os modelos não apresentam inconsistências nos seus coeficientes.

No que respeita à significância estatística dos coeficientes e considerando um nível de significância de 10%, todos os modelos apresentam, para o coeficiente mo2, um valor de $p > 0,10$, não se rejeitando a hipótese nula desse coeficiente ser igual a zero, isto é, ser nula a desutilidade associada a um ligeiro problema de mobilidade.

Comparando os critérios de informação AIC e BIC observa-se que o modelo que apresenta os menores valores para estes critérios é o GLS de efeitos variáveis (Modelo 2)

seguido dos modelos GLS censurado com coeficiente aleatório (Modelo 3) e regressão por intervalo GLS censurado em 1 ou GLS_1 (Modelo 5).

Considerando que o modelo GLS (Modelos 2 e 3) não lida adequadamente com as questões de dados censurados, optou-se pelo modelo de regressão por intervalo GLS censurado em 1 (Modelo 5) para a estimação do sistema de valores do EQ-5D-5L a partir dos dados cTTO, pelo facto de ser o modelo que melhor se adequa na modelação de dados com intervalos censurados. Na medida em que o processo iterativo é determinante para a precisão dos valores cTTO, nestas circunstâncias desconhece-se o valor exacto da observação mas sabe-se o intervalo onde recai.

Assim, tendo em conta os resultados do modelo escolhido (Modelo 5), pode calcular-se, a título de exemplo, o valor de um estado de saúde específico (e.g. 14235). Este exemplo de estado de saúde descreve um indivíduo sem problemas de mobilidade mas com problemas graves em se lavar ou vestir, problemas ligeiros em desempenhar as suas actividades habituais, com dores ou mal-estar moderados e extremamente ansioso ou deprimido. Partindo-se de um estado de saúde perfeita – ao qual se atribuiu o valor 1 – o valor ou utilidade atribuído ao estado de saúde 14235, considerando os resultados obtidos no modelo, é calculado subtraindo-se a esse valor, os valores da constante e dos coeficientes associados aos níveis de problemas de cada dimensão que caracterizam o estado, conforme explicitado na equação 10:

$$1 - 0,031 - 0,000 - 0,164 - 0,032 - 0,092 - 0,291 = 0,390$$

A análise de sensibilidade efectuada para verificar a robustez dos parâmetros obtidos nos modelos econométricos, teve por objectivo analisar o impacto da decisão de exclusão de dados associados aos estados de saúde assinalados no módulo de feedback. Neste âmbito, a análise de sensibilidade consistiu na avaliação do impacto da aplicação deste critério de exclusão, i.e., na avaliação da consistência da decisão na escolha do modelo considerando a totalidade dos dados de valoração (n=14.510) sem qualquer tipo de exclusão.

Tabela 22 - Resultados da análise de sensibilidade (modelos para dados cTTO)

Modelos Critérios	Modelo 1: OLS	Modelo 2: GLS efeitos variáveis	Modelo 3: GLS censurado	Modelo 4: Tobit censurado	Modelo 5: Intervalo GLS_1	Modelo 6: Intervalo GLS_1_0
Inconsistências	0	0	0	0	0	0
Sem significância estatística ($\alpha=10\%$)	2	1	1	3	1	1
AIC	18134,434	12952,519	16411,164	21809,051	16411,164	16876,595
BIC	18239,669	13126,918	16585,564	21975,869	16585,564	17050,994

A análise à tabela 22 permite verificar que, considerando a totalidade das 14.510 valorações de estados de saúde obtidas neste estudo, os modelos econométricos não apresentam inconsistências. No entanto, assinala-se que o coeficiente associado ao nível 2 de intensidade de problemas na dimensão mobilidade (mo2) apresenta um valor negativo (-0,002) no modelo Tobit censurado (Modelo 4).

No que respeita à significância estatística dos coeficientes e considerando um nível de significância de 10%, observa-se um maior número de ocorrências de coeficientes sem significância estatística, quando comparado com os resultados dos modelos com exclusão de dados. Uma vez mais todos os modelos apresentam para o coeficiente mo2 um valor de $p > 0,10$, não se rejeitando a hipótese nula desse coeficiente ser igual a zero. O mesmo sucede com o coeficiente ua2 para o modelo OLS (Modelo 1) e os coeficientes ua2 e ad2 para o modelo Tobit censurado (Modelo 4).

Comparando os critérios de informação AIC e BIC, observa-se, de igual forma, que o modelo que apresenta os menores valores para estes critérios é o GLS de efeitos variáveis (Modelo 2) seguido dos modelos GLS censurado com coeficiente aleatório (Modelo 3) e regressão por intervalo GLS censurado em 1 (Modelo 5).

Da análise destes critérios conclui-se que a exclusão de dados não tem impacto nas decisões tomadas relativamente à escolha do modelo econométrico para a estimação do sistema de valores do EQ-5D-5L, a partir dos dados cTTO, quando aplicado o critério de exclusão dos estados de saúde assinalados no módulo de feedback.

6.4. Modelo DCE

6.4.1. Controlo de qualidade

Numa perspectiva de controlo de qualidade e no que concerne aos padrões pouco usuais de resposta associados à tarefa DCE (e.g. sequências de valorações do tipo AAAAAAA, BBBBBBB e ABABABA), apresenta-se na tabela 23 o número obtido de ocorrências através da ferramenta QCTool.

Tabela 23 - Quantificação dos padrões de resposta pouco usuais da DCE (n=1.451)

AAAAAAA	BBBBBBB	ABABABA	BABABAB
3	1	16	9

Considerando os 1.451 participantes, apenas 29 apresentaram padrões de resposta pouco usuais, sendo que o mais frequente (n=16) foi a combinação ABABABA. Face ao número de indivíduos incluídos no estudo, considera-se que o número de respostas sequenciais pouco comuns foi muito pouco expressiva ocorrendo apenas em 2% dos participantes.

6.4.2. Análise dos dados

Nesta segunda técnica de valoração, cada um dos 1.451 participantes contribuiu com a valoração de sete pares de escolha, o que resultou num total de 10.157 valorações de estados de saúde.

No processo de elicitación, o participante demorou, em média, 54 segundos (0,9 minutos) a responder, quando perante cada par de escolhas.

Na componente DCE não se aplicou qualquer exclusão de dados na medida em que, neste estudo, a exclusão incidiu sobre os estados de saúde assinalados pelo participante no módulo de feedback que integra a componente de valoração cTTO, não influenciando portanto a tarefa de valoração DCE.

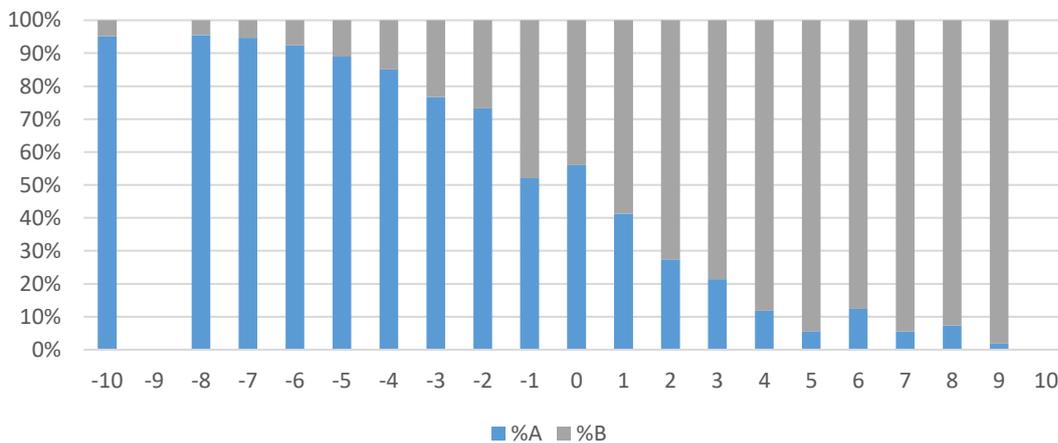
Tal como referido anteriormente, cada uma das tarefas DCE envolveu a escolha entre dois estados de saúde: Vida A e Vida B. Para cada um dos estados de saúde é possível estimar o seu índice de intensidade de problemas, somando os seus níveis. Daqui tem-se

que a intensidade pode variar entre 5 (estado de saúde representado por 11111) e 25 (estado de saúde representado por 55555). Se se considerar um par de escolhas com estas duas opções, a diferença entre a intensidade será de $5 - 25 = -20$.

Neste estudo, a diferença de intensidade entre os estados de saúde dos pares de escolha variou entre -10 e 9.

O gráfico 16 apresenta a percentagem de respondentes que escolheu a Vida A comparada com a diferença de intensidade de problemas do estado de saúde da Vida A e do estado de saúde da Vida B.

Gráfico 16 - Percentagem de participantes que escolheu a Vida A ou a Vida B na tarefa DCE comparada com a intensidade relativa de problemas de A e B



A análise ao gráfico 16 permite verificar que existe variabilidade nas respostas dos participantes e que estas não foram condicionadas pela ordem das alternativas.

A tabela 24 apresenta os resultados dos modelos utilizados na modelação de dados DCE e a avaliação do desempenho de cada um deles, de acordo com os critérios previamente definidos.¹⁴

¹⁴ Para cada um dos coeficientes estimados foram calculados também intervalos de confiança e testada a presença de heteroscedasticidade nos modelos.

Tabela 24 - Resultados e avaliação dos modelos Logit condicional, Logit e Probit (n=10.157)

EQ-5D	Modelo 7: Probit			Modelo 8: Logit			Modelo 9: Logit condicional		
	β	Erro-Padrão	p	β	Erro-Padrão	p	β	Erro-Padrão	p
mo2	0,226	0,030	0,000	0,384	0,049	0,000	0,384	0,049	0,000
mo3	0,341	0,034	0,000	0,598	0,058	0,000	0,598	0,058	0,000
mo4	0,641	0,035	0,000	1,104	0,060	0,000	1,104	0,060	0,000
mo5	1,247	0,039	0,000	2,169	0,070	0,000	2,169	0,070	0,000
cp2	0,128	0,032	0,000	0,245	0,055	0,000	0,245	0,055	0,000
cp3	0,210	0,035	0,000	0,362	0,060	0,000	0,362	0,060	0,000
cp4	0,500	0,035	0,000	0,871	0,060	0,000	0,871	0,060	0,000
cp5	1,006	0,035	0,000	1,747	0,062	0,000	1,747	0,062	0,000
ah2	0,170	0,030	0,000	0,280	0,052	0,000	0,280	0,052	0,000
ah3	0,182	0,034	0,000	0,313	0,058	0,000	0,313	0,058	0,000
ah4	0,419	0,034	0,000	0,720	0,059	0,000	0,720	0,059	0,000
ah5	0,843	0,035	0,000	1,490	0,062	0,000	1,490	0,062	0,000
dm2	0,194	0,032	0,000	0,355	0,055	0,000	0,355	0,055	0,000
dm3	0,376	0,034	0,000	0,662	0,059	0,000	0,662	0,059	0,000
dm4	0,876	0,036	0,000	1,538	0,063	0,000	1,538	0,063	0,000
dm5	1,367	0,038	0,000	2,413	0,071	0,000	2,413	0,071	0,000
ad2	0,057	0,034	0,090	0,133	0,058	0,021	0,133	0,058	0,021
ad3	0,259	0,033	0,000	0,463	0,058	0,000	0,463	0,058	0,000
ad4	0,624	0,036	0,000	1,129	0,065	0,000	1,129	0,065	0,000
ad5	0,870	0,037	0,000	1,526	0,066	0,000	1,526	0,066	0,000
Inconsistências	0			0			0		
Sem significância estatística ($\alpha=10\%$)	0			0			0		
AIC	9.584,165			9,557,849			9.557,849		
BIC	9.728,684			9.702,367			9.716,230		

Nota: mo – mobilidade; cp – cuidados pessoais; ah – actividades habituais; dm – dor/mal-estar; ad – ansiedade/depressão

Na verdade, todos os modelos avaliados apresentam parâmetros consistentes e com significância estatística ($p < 0,10$). A diferença entre os modelos obtém-se através dos critérios de informação AIC e BIC, através dos quais se conclui que o modelo que apresenta um melhor ajustamento é o Logit (Modelo 8).

Note-se que, pela natureza subjacente a este tipo de modelos econométricos, os parâmetros resultantes não correspondem aos coeficientes associados a cada nível de intensidade de problemas, dentro de cada dimensão, não podendo, por isso, ser directamente comparados aos coeficientes obtidos nos modelos econométricos utilizados

para os dados cTTO. No entanto, os parâmetros obtidos podem ser sujeitos à avaliação de desempenho de acordo com os critérios definidos para efeitos de avaliação dos modelos. De forma a obter valores passíveis de comparação com os alcançados no modelo cTTO, utiliza-se a função logística inversa para transformar os coeficientes do modelo Logit.

No contexto da equação 10 e considerando o exemplo anteriormente apresentado relativo ao valor atribuído ao estado de saúde 14235, bem como os coeficientes estimados pelo modelo Logit (Modelo 8), o valor ou utilidade atribuído ao estado de saúde obtém-se fazendo a seguinte transformação:

$$1 - 0,000 - 0,000 - 0,871 - 0,280 - 0,662 - 1,526 = -2,339$$

$$P(-2,339) = \frac{\exp(-2,339)}{(1 + \exp(-2,339))} = 0,088$$

Recorda-se que, neste estudo, os valores dos estados de saúde com recurso à técnica DCE não foram estimados numa escala de utilidade. Por este motivo, os resultados obtidos neste estudo com recurso a esta técnica não podem ser usados de forma independente para estimar o valor dos estados de saúde (Rowen et al. 2015; Stolk et al. 2010).

6.5. Modelo híbrido

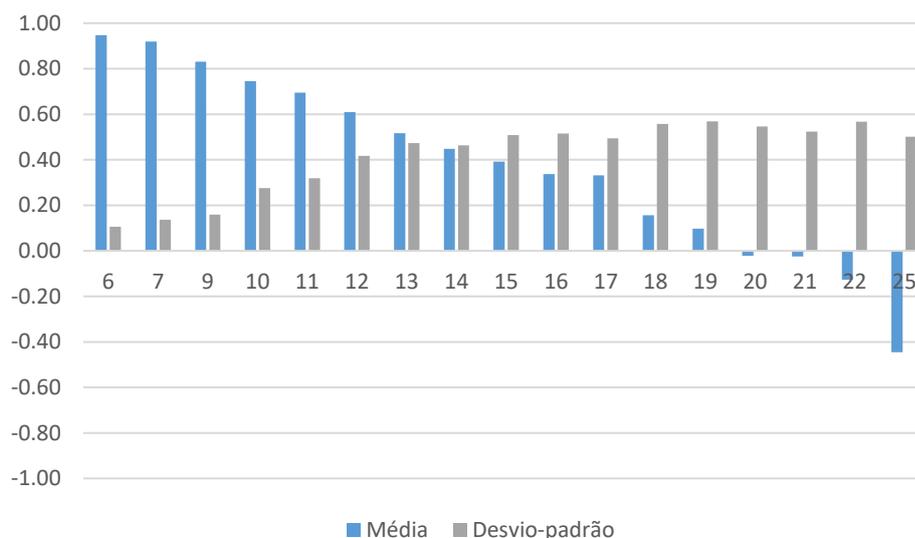
Tal como referido anteriormente, o modelo híbrido permite a combinação dos dados contínuos da tarefa cTTO com os dados dicotómicos da tarefa DCE numa função única de máxima verosimilhança. A sua aplicação no Stata é realizada através do comando Hyreg (Ramos-Goñi et al. 2016).

Na medida em que o modelo híbrido incorpora os dados decorrentes dos modelos cTTO e DCE para os quais foi anteriormente apresentado o controlo de qualidade, este subcapítulo respeita apenas à análise dos dados.

No que concerne à exclusão de dados, na modelação híbrida aplicou-se uma vez mais a exclusão para os dados cTTO dos estados de saúde assinalados no módulo de feedback.

Considerando a dimensão da amostra presente neste estudo, é de esperar que exista heterogeneidade na valoração dos estados de saúde. De facto, o gráfico 17 permite observar que, à medida que o índice de intensidade dos problemas associados a cada estado de saúde aumenta, o valor médio dos estados de saúde em análise, obtido com a técnica de valoração cTTO diminui, alcançando resultados médios negativos a partir de pontuações superiores a 20. Por outro lado, verifica-se que o desvio-padrão aumenta particularmente após pontuações superiores a 10.

Gráfico 17 - Valores médios e desvio-padrão por pontuação obtida nos níveis de problemas dos estados de saúde valorados pelo cTTO



Uma opção para responder a esta situação de variabilidade heterogénea, transversal à valoração dos estados de saúde, é a utilização de modelos econométricos heteroscedásticos.

A aplicação do modelo híbrido heteroscedástico permitiu obter parâmetros consistentes e estatisticamente significativos, exceptuando o valor associado à constante que apresenta um valor de $p > 0,10$, não se rejeitando a hipótese nula da constante ser igual a zero (Tabela 25). De acordo com os autores do modelo híbrido, não existem até ao

momento medidas estatísticas que permitam comparar o desempenho do modelo híbrido com outros. Os critérios de informação básicos como o AIC e o BIC não se aplicam na medida em que a equação da verosimilhança logarítmica (ou *log likelihood*) de um modelo híbrido constitui a soma dos dois modelos sendo, por definição, superior à de um modelo para dados contínuos ou de um modelo para dados dicotómicos. Outras medidas, como o erro médio quadrado ou o erro médio absoluto, não poderão ser utilizadas nos modelos híbridos, por ausência de evidência actual que suporte a sua utilização.

Tabela 25 - Resultados do modelo híbrido heteroscedástico (Modelo 10)

EQ-5D-5L	β	Erro-Padrão	p	[Intervalo de Confiança 95%]	
mo2	0,049	0,005	0,000	0,039	0,059
mo3	0,090	0,007	0,000	0,076	0,104
mo4	0,176	0,007	0,000	0,162	0,190
mo5	0,333	0,007	0,000	0,319	0,346
cp2	0,050	0,005	0,000	0,040	0,061
cp3	0,071	0,007	0,000	0,057	0,085
cp4	0,153	0,007	0,000	0,139	0,167
cp5	0,278	0,007	0,000	0,265	0,290
ah2	0,045	0,005	0,000	0,035	0,055
ah3	0,065	0,007	0,000	0,051	0,078
ah4	0,132	0,007	0,000	0,119	0,146
ah5	0,249	0,007	0,000	0,236	0,262
dm2	0,043	0,005	0,000	0,033	0,052
dm3	0,100	0,007	0,000	0,086	0,115
dm4	0,243	0,007	0,000	0,229	0,257
dm5	0,382	0,008	0,000	0,368	0,397
ad2	0,038	0,005	0,000	0,028	0,048
ad3	0,087	0,007	0,000	0,073	0,100
ad4	0,207	0,007	0,000	0,194	0,220
ad5	0,268	0,007	0,000	0,254	0,281
constante	0,000	0,003	0,923	-0,007	0,006
13.391 observações contínuas (CTTO)					
10.157 observações dicotómicas (DCE)					

Nota: mo – mobilidade; cp – cuidados pessoais; ah – actividades habituais; dm – dor/mal-estar; ad – ansiedade/depressão

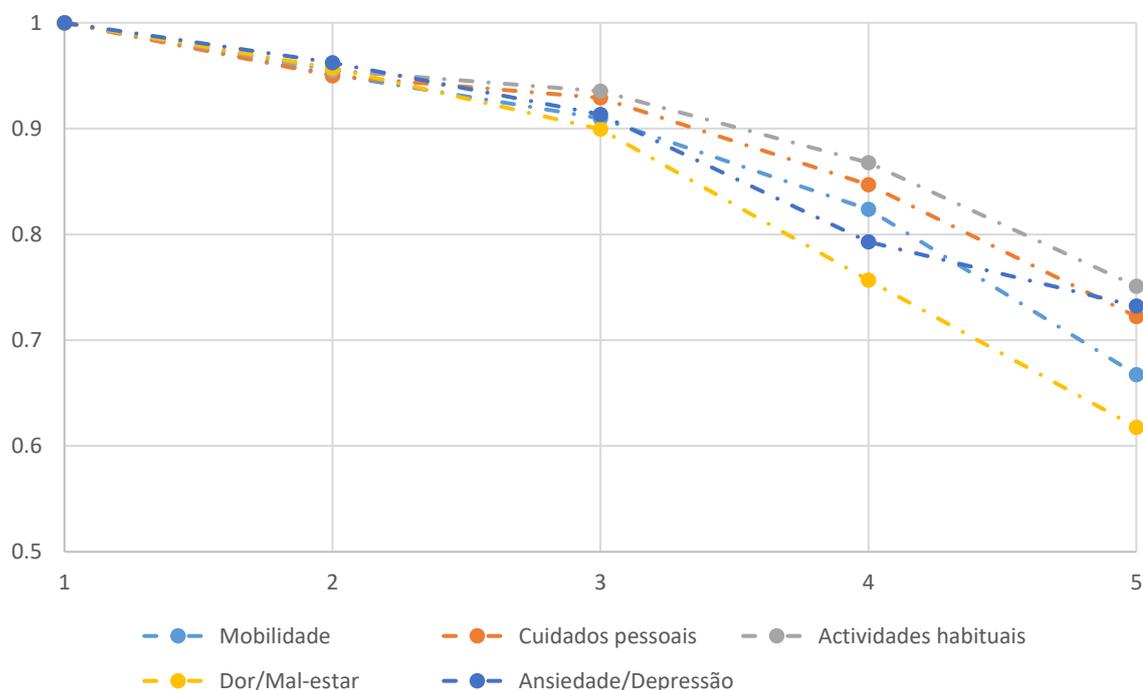
Considerando que os coeficientes obtidos no modelo traduzem a desutilidade associada ao aumento do nível de problemas, dentro de cada dimensão do EQ-5D-5L, se se partir de um estado de saúde perfeita, ao qual se associa o valor de 1, é possível subtrair o valor de cada coeficiente e, assim, obter a valoração do estado de saúde.

Partindo uma vez mais do exemplo previamente utilizado relativo ao valor do estado de saúde 14235 e no contexto da equação 10, tem-se:

$$1 - 0,000 - 0,000 - 0,153 - 0,045 - 0,100 - 0,268 = 0,434$$

O gráfico 18 pretende mostrar o valor dos estados de saúde, dentro de cada dimensão, à medida que aumenta o nível de intensidade dos problemas. Tratando-se de resultados numa escala ordinal, o tracejado do gráfico existe apenas para facilitar a sua leitura. Uma vez mais observa-se que os resultados não apresentam inconsistências e que, de uma perspectiva societária, as dimensões dor/mal-estar e mobilidade são as piores valoradas no nível 5 de intensidade de problemas o que vai ao encontro das dimensões do sistema descritivo do EQ-5D-5L consideradas pelos participantes como mais relevantes, durante a execução das tarefas de valoração cTTO e DCE.

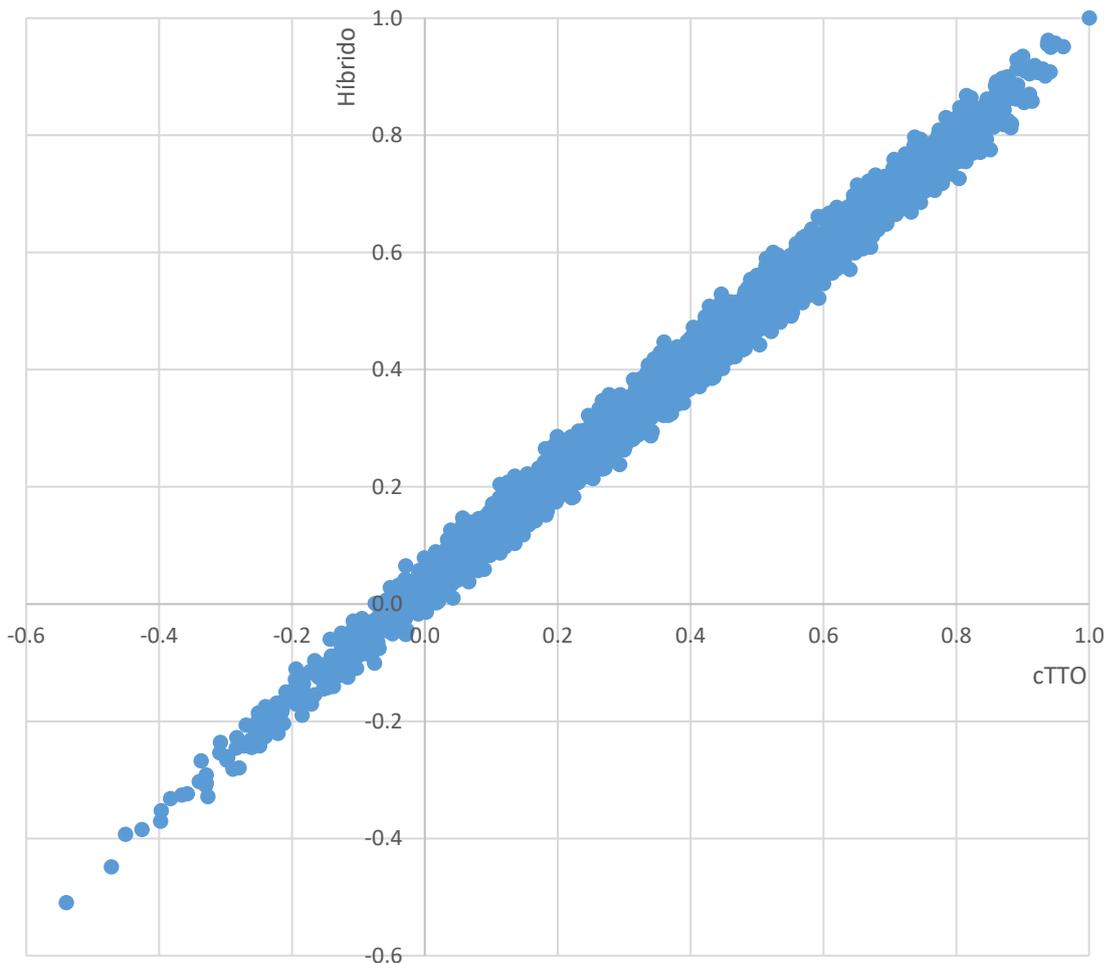
Gráfico 18 - Valor dos estados de saúde por nível e dimensão do EQ-5D-5L



O gráfico permite ainda verificar que níveis caracterizados por mais problemas estão sempre associados a menor valor. No entanto, essa degradação não é tão acentuada

na dimensão ansiedade/depressão. Quando se analisam as valorações obtidas para os 3.125 estados de saúde gerados pelo EQ-5D-5L considerando o modelo de regressão por intervalo GLS censurado em 1 (Modelo 5), que integra os dados da tarefa cTTO, e o modelo híbrido heteroscedástico (Modelo 10), que combina os dados contínuos da tarefa cTTO com os dados dicotómicos da tarefa DCE, o diagrama de dispersão permite observar que existe correlação linear, na medida em que é possível ajustar uma recta à nuvem de pontos obtida (Gráfico 19). A intensidade da associação linear existente é positiva e forte, considerado que o coeficiente de correlação linear de Pearson atinge um valor muito próximo de 1 ($r = 0,995$) o que evidencia a forte relação que existe entre as valorações geradas pelos dois modelos.

Gráfico 19 - Comparação dos valores estimados para os 3.125 estados de saúde gerados pelo EQ-5D-5L com recurso ao modelo de regressão por intervalo GLS censurado em 1 e ao modelo híbrido heteroscedástico



Para além do modelo híbrido heteroscedástico testou-se, ainda neste estudo, a versão híbrida censurada e heteroscedástica, não tendo, no entanto, sido alcançada convergência do modelo pelo que não se apresentam os resultados obtidos.

6.6. Comparação dos sistemas de valores, cTTO e híbrido, obtidos para Portugal com outros sistemas de valores internacionais

Tal como referido anteriormente, as características demográficas e os aspectos culturais de cada país podem influenciar a valoração da saúde (Tsuchiya et al. 2002; Jelsma et al. 2003; Johnson et al. 2005; Yusof et al. 2012). Esse é o motivo pelo qual se torna relevante a condução de estudos nacionais de valoração. Nesse âmbito e considerando os estudos de valoração do EQ-5D-5L já finalizados e cujo sistema de valores foi publicado, apresenta-se nos gráficos seguintes uma comparação entre os sistemas de valores do EQ-5D-5L obtidos nesta tese para Portugal, tendo por base os modelos cTTO e híbrido escolhidos, e os sistemas de valores do EQ-5D-5L obtidos para a Holanda, Uruguai, Inglaterra e Japão (Versteegh et al. 2016; Augustovski et al. 2016; Devlin et al. 2016; Ikeda et al. 2015).

Os gráficos 20 e 21 apresentam, respectivamente, a comparação entre o sistema de valores do EQ-5D-5L obtido para Portugal, com recurso aos dados CTTO, e o sistema de valores do EQ-5D-5L para a Holanda e Uruguai, recorrendo à mesma técnica de valoração.

Gráfico 20 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Holanda (modelo cTTO)

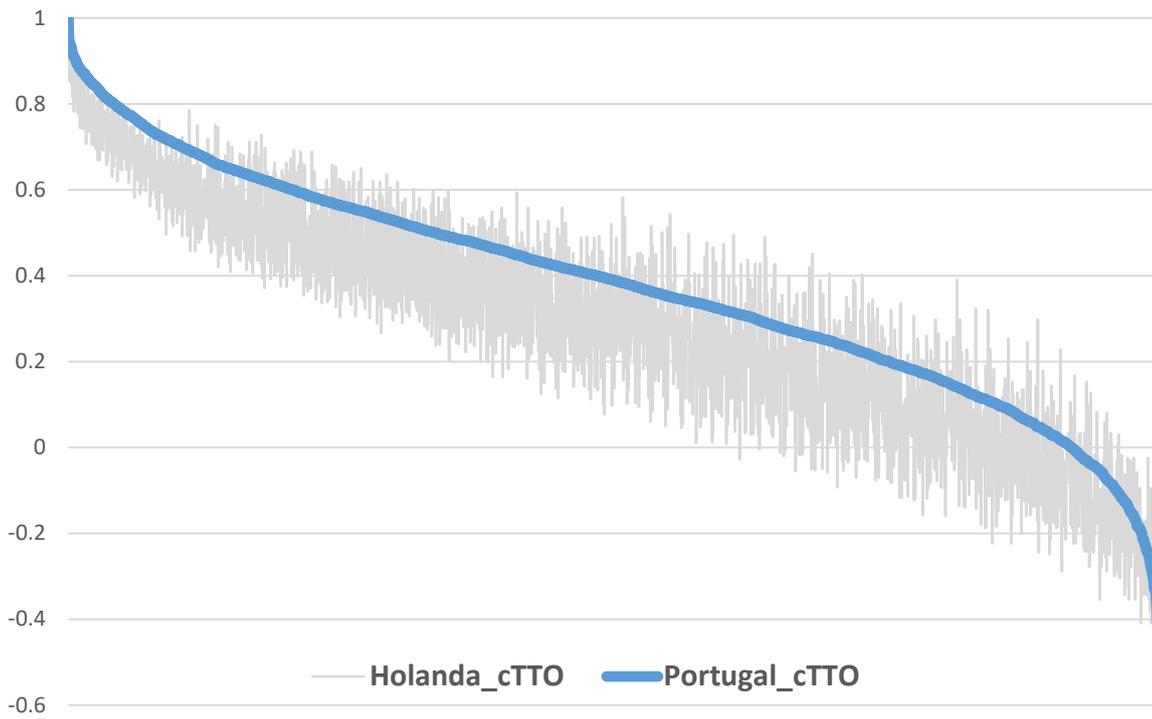
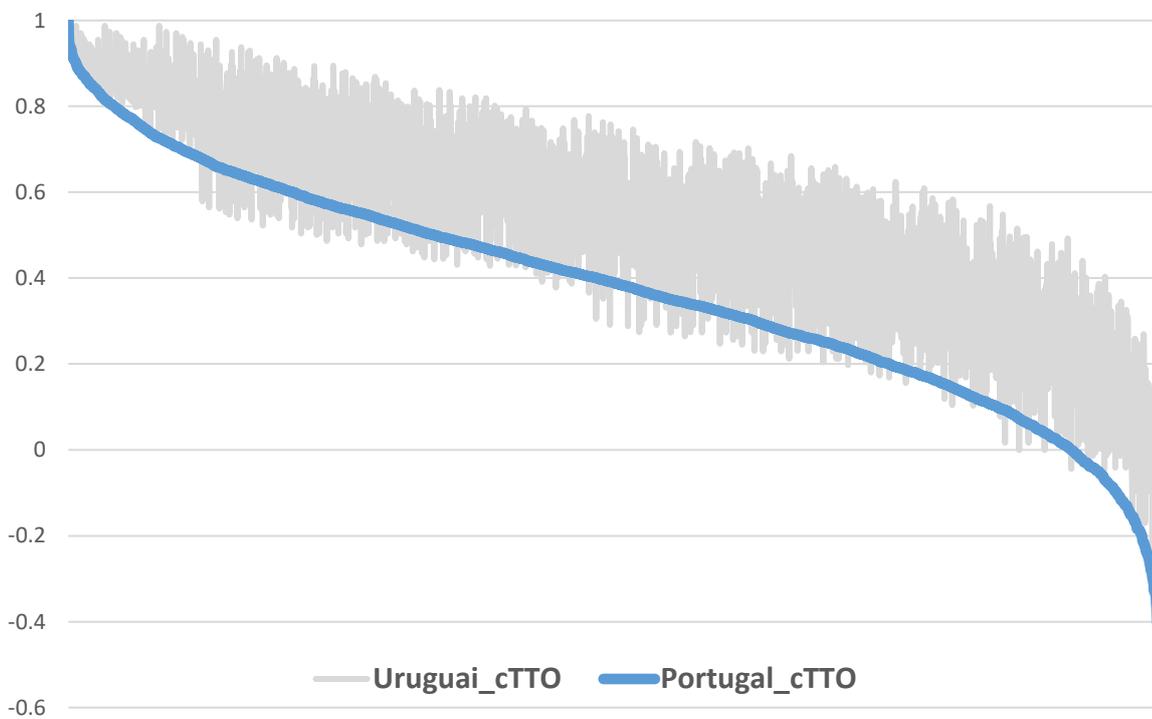


Gráfico 21 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Uruguai (modelo cTTO)



Tal como expectável, verifica-se que o sistema de valores português é diferente dos sistemas de valores obtidos para a Holanda e Uruguai. No entanto e considerando o modelo cTTO, o maior nível de semelhança ocorre com o sistema de valores holandês, obtendo-se um coeficiente de correlação linear de Pearson de 0,933. Já no que concerne ao sistema de valores do Uruguai, quando comparado com o sistema de valores português, o valor do coeficiente de correlação linear de Pearson é de 0,900.

Os gráficos 22 e 23 efectuam a mesma comparação mas agora entre o sistema de valores obtido para Portugal tendo por base o modelo híbrido e os sistemas de valores publicados para Inglaterra e Japão, que tiveram também por base a combinação de dados contínuos (cTTO) e dados dicotómicos (DCE).

Gráfico 22 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Inglaterra (modelo híbrido)

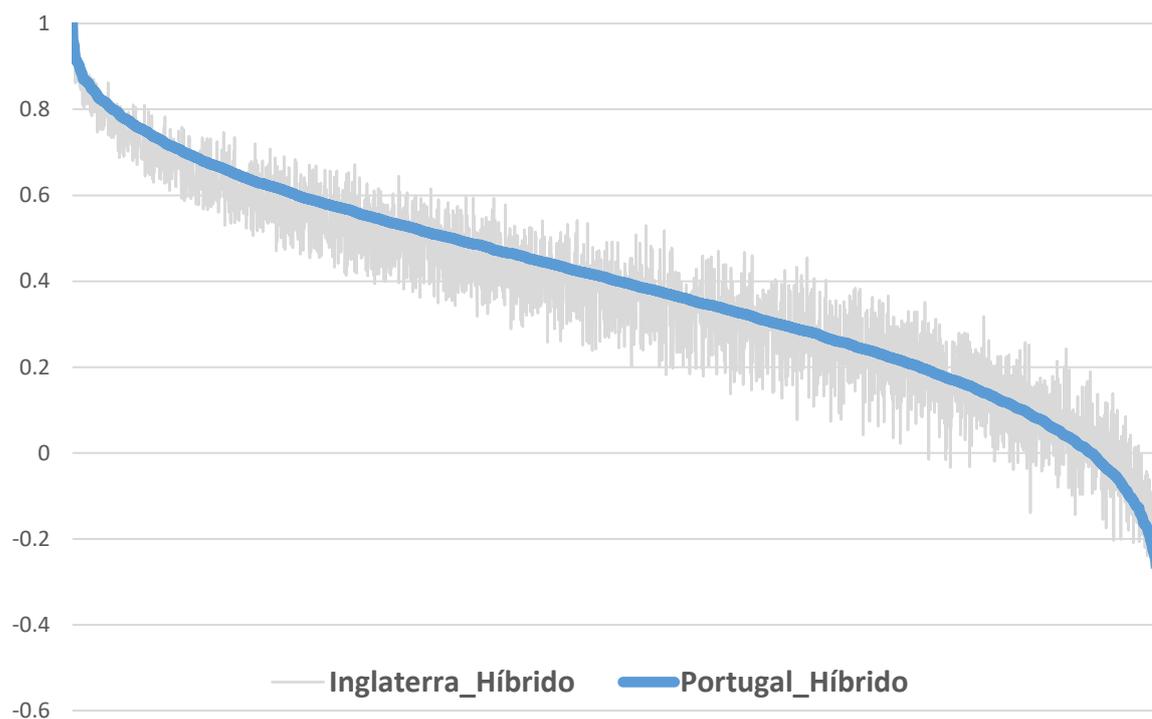
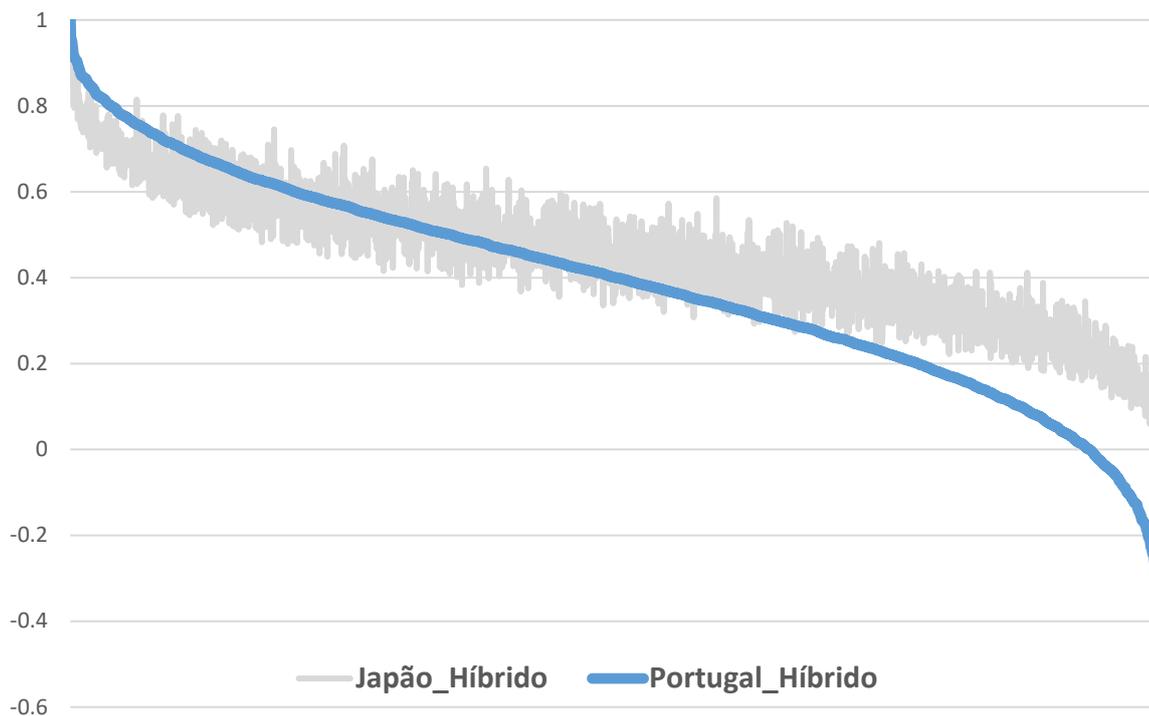


Gráfico 23 - Comparação do sistema de valores do EQ-5D-5L obtido para Portugal e Japão (modelo híbrido)



Não obstante os sistemas de valores serem diferentes, é interessante observar a semelhança que existe entre o sistema de valores português e o de Inglaterra quando utilizados dados híbridos. Neste caso concreto, o valor do coeficiente de correlação linear de Pearson é muito próximo de 1 ($r = 0,962$). Considerando o sistema de valores publicado para o Japão, esse valor é de 0,951.

7. DISCUSSÃO

O principal objectivo deste estudo foi estimar o sistema de valores do EQ-5D-5L para a população portuguesa, com recurso a duas técnicas de elicitación de preferências: cTTO e DCE. A determinação do valor que os 1.451 participantes, representativos da população portuguesa, associam aos 3.125 estados de saúde gerados pelo sistema descritivo do EQ-5D-5L possibilita a utilização deste sistema de valores em estudos de avaliação económica em saúde, tendo por objectivo o apoio à decisão sobre diferentes intervenções de saúde, num contexto nacional de priorização e alocação de recursos limitados.

Os resultados deste estudo sugerem que os problemas de saúde associados às dimensões dor/mal-estar e mobilidade são os mais relevantes, considerando as preferências da população portuguesa. Já os problemas inerentes à realização de actividades habituais (trabalho, estudo, actividades domésticas, actividades em família ou de lazer) e os associados à ansiedade/depressão revelaram-se como os menos marcantes. A importância destes resultados reside no facto de poderem ser utilizados na avaliação de diferentes opções, priorizando a utilização dos recursos e, ao mesmo tempo, respeitando as preferências da população portuguesa.

A nível internacional são ainda em número reduzido os estudos publicados de valoração do EQ-5D-5L. Neste âmbito, identificam-se os estudos realizados no Canadá (Xie et al. 2016), Inglaterra (Devlin et al. 2016), Holanda (Versteegh et al. 2016), Uruguai (Augustovski et al. 2016), Japão (Ikeda et al. 2015) e Coreia (Kim et al. 2016). Não obstante ter sido publicado, em 2014, um primeiro estudo relativo ao sistema de valores espanhol, os autores não recomendam a sua utilização, encontrando-se a rever a sua análise (Ramos-Goñi et al. 2014).

No que concerne à utilização de dados híbridos (cTTO e DCE) são ainda mais escassos os exemplos da combinação destes dois tipos de informação para a estimação do sistema de valores. Nesta data, identificam-se os estudos efectuados em Inglaterra e no Japão (Devlin et al. 2016; Feng et al. 2016; Ikeda et al. 2015). Apesar de alguns dos restantes

estudos de valoração mencionados terem contemplado a utilização de mais do que uma técnica de eliciação de preferências (cTTO e DCE), a modelação dos dados efectuou-se utilizando separadamente dados contínuos e dicotómicos, tendo o modelo econométrico para estimação do sistema de valores recaído na opção conservadora de um modelo exclusivo para dados contínuos (cTTO). A comparação efectuada entre os sistemas de valores do EQ-5D-5L, obtidos nesta tese para Portugal, considerando os modelos cTTO e híbrido, e os sistemas de valores conhecidos dos estudos identificados anteriormente mostra que, apesar das semelhanças encontradas com os resultados do estudo holandês, que utilizou um modelo cTTO, e com os resultados do estudo inglês, que recorreu a um modelo híbrido, ainda assim verificam-se diferenças possivelmente justificadas pelas características únicas associadas à população de cada país, o que justifica a importância de se obter um sistema de valores nacional para utilização em estudos de avaliação económica.

O grupo coordenador nacional implementou este estudo seguindo o código de conduta e o protocolo de valoração EQ-VT criados pelo grupo EuroQol. O protocolo EQ-VT foi desenvolvido com vista à normalização dos estudos de valoração efectuados nos diversos países, permitindo a obtenção de resultados mais válidos e comparáveis a nível internacional.

O EQ-5D-5L constitui a versão mais recente do instrumento de medição em saúde EQ-5D, desenvolvido pelo grupo EuroQol. A diferença relativamente à anterior versão (EQ-5D-3L) reside no aumento do número de níveis, de três para cinco, dentro de cada uma das cinco dimensões (Herdman et al. 2011). Este aumento no número de níveis do sistema descritivo permite uma melhor distinção dos estados de saúde e, no limite, uma maior discriminação na valoração dos mesmos (M. Janssen et al. 2013).

No entanto, a metodologia seguida neste estudo suscita um conjunto de argumentos passíveis de discussão, origina questões ainda por responder e sugere áreas para investigação futura.

Em primeiro lugar, surge um aspecto directamente relacionado com a implementação do estudo e que se relaciona com a organização da recolha de dados. O estudo contemplou duas equipas de entrevistadores, uma supervisionada pela empresa de

sondagens e outra dirigida pelo grupo coordenador nacional. A principal diferença nestes dois grupos residiu na duração da recolha dos dados. O grupo de 14 entrevistadores sob a supervisão da empresa de sondagens demorou nove meses para realizar 750 entrevistas, enquanto que o grupo de 14 entrevistadores sob a responsabilidade do grupo coordenador nacional efectuou 701 entrevistas em um mês e meio. Não obstante o apoio e acompanhamento disponibilizados pela equipa nacional de investigadores ter sido igual para ambos os grupos, a diferença no desempenho poderá relacionar-se com o facto de a maioria dos entrevistadores cujo centro de recrutamento foi Coimbra, por ser estudante de ciclos de ensino superior e ter tido eventualmente experiência prévia para com trabalhos de investigação, sentir um maior empatia para com o estudo. O mesmo já não sucedeu com o grupo recrutado em Lisboa que poderá ter tido maior dificuldade em identificar-se com o estudo. A excepção surge com a equipa de dois entrevistadores de Évora, responsável pela execução de aproximadamente 33% das entrevistas, que manteve um trabalho de campo constante ao longo de quase todo o período de recolha de dados (nove meses). Esta diferença na identificação para com os objectivos do estudo e no desempenho obriga à reflexão sobre recrutamentos futuros, designadamente, em estudos de valoração que aplicam técnicas morosas e complexas, quer para o entrevistador quer para o participante. A discussão tida com outras equipas internacionais que conduziram estudos semelhantes conduziu a conclusões similares. Os grupos coordenadores de outros estudos concluíram que o recrutamento em contexto universitário e a supervisão directa, sem intermediários, conduz a um resultado mais favorável quer a nível da qualidade do desempenho quer a nível da duração da recolha de dados (comunicação pessoal).

Outra questão surge a nível da população alvo deste estudo. Tal como referido anteriormente, o sistema de valores estimado reflecte as preferências da população portuguesa em geral podendo ser utilizado em processos de tomada de decisão em saúde. A pergunta que surge é se o mesmo sistema de valores poderá ser utilizado para grupos de doentes com patologias ou para grupos de pessoas com características específicas, como crianças, idosos ou indivíduos de culturas diferentes. Num contexto de decisão clínica individual faz sentido a abordagem dos indivíduos afectados por essas circunstâncias (e.g. avaliação de indivíduos submetidos a técnicas cirúrgicas inovadoras face a técnicas

convencionais). Neste âmbito, o sistema de valores agora obtido pode ser questionado. No entanto, considerando um contexto abrangente de definição de prioridades em saúde e de alocação de recursos escassos, com impacto na generalidade da população e para os quais essa mesma população contribui com os seus impostos, é razoável assumir as preferências de uma amostra representativa da população como forma de estimar o valor de cada estado de saúde.

No que concerne às crianças e adolescentes, a situação assume um carácter mais particular. Para tal, o grupo EuroQol desenvolveu esforços na construção de uma versão pediátrica do instrumento de medição em saúde EQ-5D, designado por EQ-5D-Y e destinados a crianças/adolescentes dos 8 aos 15 anos de idade. A principal diferença relativamente ao EQ-5D-5L reside na forma como o sistema descritivo é redigido, adaptando a descrição dos níveis de problemas em cada dimensão para uma linguagem perceptível a este tipo particular de destinatários. Os primeiros trabalhos publicados relativamente à utilização do EQ-5D-Y surgiram em 2010 (Wille et al. 2010; Ravens-Sieberer et al. 2010). Em 2015, este instrumento encontrava-se traduzido para 40 línguas. Os primeiros passos para traduzir e adaptar o sistema descritivo do EQ-5D-Y para a versão portuguesa foram iniciados pelo CEISUC, no final de 2016. Não obstante a inexistência de dificuldades associadas à utilização do sistema descritivo, mesmo quando o alvo são crianças mais jovens, situação em que um adulto na qualidade de pai/mãe ou noutra poderá ser utilizado como *proxy*, até ao momento actual não foi possível estimar um sistema de valores associado ao EQ-5D-Y. Tal deve-se ao facto de subsistirem dúvidas por parte da comunidade científica sobre se se deve elicitar as preferências de crianças e jovens e, se sim, quem deve ser o alvo da elicitação das preferências (criança, pai/mãe, população em geral). As questões complexas não restam por aqui recaindo ainda sobre a técnica de valoração (VAS, DCE, TTO ou outra) mais adequada para obter o sistema de valores e qual a base científica para o fazer.

O trabalho que decorre desta tese utiliza duas técnicas de valoração de estados de saúde: uma técnica mais conhecida e conservadora – cTTO – e outra mais inovadora na área de avaliação económica em saúde – DCE. As especificações possíveis que podem ser

seguidas no desenho de ambas as técnicas para contextualizar o processo de escolha podem ter impacto na valoração dos estados de saúde.

No que concerne ao cTTO, esta técnica permite manter a consistência com os estudos de valoração anteriores com recurso ao EQ-5D-3L, na medida em que apenas a valoração associada aos estados de saúde considerados piores que morte segue um método distinto – variante *lead-time* TTO – de forma a evitar o redimensionamento discricionário da escala de valores, tal como sucedia com o TTO convencional ao permitir uma escolha arbitrária de tempo para troca (B. Janssen et al. 2013). Não obstante esta abordagem não ser isenta de críticas, tal como referido pelos mesmos autores e explicitado no subcapítulo 3.3 desta tese, ainda assim, o cTTO constitui uma técnica válida e consistente na eliciação de valores negativos. Apesar do cTTO incorporar o conceito de sacrifício ou troca entre qualidade e quantidade de vida, permitindo ancorar valores numa escala latente de utilidade e tornando-o, assim, uma escolha de eleição como técnica de eliciação de preferências, constitui uma técnica complexa, com maior esforço cognitivo associado, que resulta numa aplicação dispendiosa e consumidora de tempo. Por este motivo, esta tarefa mantém a necessidade de recorrer a entrevistas presenciais, obrigando à dependência de entrevistadores para a obtenção de dados. Neste contexto e de forma a evitar o enviesamento possível associado aos entrevistadores (Ramos-Goñi et al. 2017), o treino e a monitorização constante destes tornam-se cruciais e, de facto, constituíram uma componente significativa da implementação deste estudo e da supervisão da recolha de dados.

Outro aspecto metodológico relevante nesta técnica de valoração consiste na ordem pela qual as dimensões do EQ-5D-5L são apresentadas na descrição de cada estado de saúde a valorar. Mulhern, Shah, Janssen, Longworth e Ibbotson (2016) sugerem que a ordem das dimensões pode ter algum impacto na magnitude dos coeficientes associados a cada dimensão e nível de problemas, com consequente efeito na valoração obtida para os estados de saúde, e suscitam questões sobre se futuros estudos de valoração deverão apresentar aleatoriamente as dimensões de saúde. O estudo efectuado por Rand-Hendriksen & Augestad (2012) corrobora esta necessidade ao evidenciar que a importância relativa das primeiras duas dimensões do EQ-5D-5L (mobilidade e cuidados pessoais) difere

de acordo com a técnica de valoração utilizada. No entanto, o estudo efectuado por Tsuchiya, Mulhern, Bansback e Hole (2014), que comparou ordens distintas na apresentação das dimensões na tarefa de valoração DCE, já não encontrou um impacto sistemático nos coeficientes obtidos. De facto, desde que as dimensões sejam apresentadas de uma só vez, a ordem pela qual surgem pode ser mais irrelevante.

Analisando ainda a técnica cTTO, esta envolve uma escolha entre um número fixo de anos num estado de saúde subóptimo e um número variável de anos num estado de saúde definido como comparador que pode ser descrito de formas distintas. Neste estudo, atribuiu-se a designação de “saúde perfeita” ao estado de saúde que serviu de comparação. Shah, Mulhern, Longworth e Janssen (2016) analisaram a diferença nos valores obtidos quando o comparador foi designado como “saúde perfeita”, tal como sucede nos estudos de valoração mais recentes (Ramos-Goñi et al. 2017), e quando foi classificado como o melhor estado de saúde que o sistema descritivo do EQ-5D-5L permite, i.e. 11111, tal como ocorreu em estudos de valoração anteriores (EuroQol Group 2005; Dolan 1997). Os resultados obtidos sugerem que os valores elicitados não foram significativamente afectados utilizando as duas designações distintas para o comparador, sendo, por isso, indiferente a classificação utilizada.

No que diz respeito à DCE, esta técnica é actualmente considerada como promissora nos estudos de valoração (Antunes et al. 2017). As principais vantagens identificadas decorrem do menor fardo cognitivo associado à sua execução, ao facto dos valores gerados não serem afectados por preferências temporais (tal como pode suceder com o cTTO), pela ausência de processos iterativos, baixo custo de implementação associado e possibilidade de execução da tarefa *online*. De acordo com a apresentação efectuada no 2nd *EuroQol Academy Meeting*, em Março de 2017, no que concerne à despesa nos estudos de valoração, o grupo EuroQol associou um custo de 160\$ por respondente no estudo australiano com recurso ao protocolo EQ-VT e de apenas 15\$ por participante quando se considerou unicamente a tarefa DCE. A utilização isolada de dados oriundos desta técnica na modelação fomenta grandes expectativas nos estudos de valoração de estados de saúde. Um aspecto particularmente interessante é o facto dos valores derivados apenas respeitarem à atractividade do estado de saúde em análise, não

sendo expressos como troca entre qualidade e quantidade de vida, minimizando, assim, o enviesamento que estes factores externos podem trazer na valoração do estado de saúde (Stolk et al. 2010).

No entanto, a técnica padece ainda de algumas limitações. Em primeiro lugar, identifica-se a dificuldade por parte da comunidade científica em aceitá-la como igual perante outras técnicas de valoração mais reconhecidas como o TTO ou o SG. No entanto, a limitação mais significativa respeita à dificuldade em ancorar os coeficientes produzidos na escala latente de utilidade. Vários métodos têm sido testados de forma a ancorar os valores modelados de utilidade na escala do QALY que varia entre 0, estado equivalente a morte, e 1, estado equivalente a saúde perfeita (Stolk et al. 2010; Rowen et al. 2015). Um dos métodos recorre aos valores obtidos para o melhor e para o pior estado de saúde definidos pelo sistema descritivo, com recurso, por exemplo, ao TTO, de forma a adaptar e ancorar os valores DCE na escala latente de utilidade. Este método é, no entanto, questionável porque faz depender a modelação DCE em técnicas de valoração existentes ao invés de promover a utilização desta técnica de forma individual (Stolk et al. 2010). Outra possibilidade é desenhar a DCE de forma a que cada estado de saúde seja sempre comparado com a alternativa de morte, isto é, de modo a perceber se o indivíduo considera que o estado de saúde em análise é melhor ou pior que morte. O valor da diferença entre esses estados de saúde e o estado de morte seria, neste caso, estimado a partir das probabilidades observadas entre esses estados de saúde e o estado de morte. Quanto ao modelo de regressão, este assume características aditivas e inclui uma variável *dummy* para o estado de morte. Os coeficientes da regressão são normalizados na escala de utilidade dividindo o coeficiente obtido para cada nível e dimensão pelo coeficiente da variável *dummy* associada à morte. Este método pode, no entanto, acarretar problemas quando os participantes não consideram que os estados de saúde em apreço são piores que a alternativa de morte utilizada para comparação. Neste caso, o modelo não captaria a distribuição dos erros e produziria estimadores enviesados (Rowen et al. 2015). Adicionalmente, de acordo com a teoria de utilidade aleatória, as respostas dos indivíduos que se comportam como *non-traders*, i.e., que consideram todos os estados de saúde dignos de ser vividos não abdicando de tempo de vida, traduzem um valor infinito

relativamente à diferença entre os estados de saúde e o estado de morte. Isto impede uma representação exacta das preferências, causa problemas na estimação dos valores e faz depender a modelação DCE na proporção de indivíduos que exibem estas características. Um outro método que tem sido testado é a inclusão da duração (tempo) como um atributo no desenho da DCE (Bansback et al. 2014). Enquanto que os dois primeiros métodos permitem ancorar os coeficientes na escala de utilidade sem necessidade de mais informação, o terceiro método, ao incluir mais um atributo, torna a DCE mais complexa e implica que um maior conjunto de informação seja analisado num único par de escolha. Como todos os métodos apresentam limitações para ancorar os valores na escala de utilidade, é complexo concluir qual a melhor estratégia a seguir.

Neste estudo, a DCE não utilizou qualquer um destes métodos elencados, impossibilitando assim a sua utilização como técnica individual, na medida em que não ancora os valores gerados na escala de utilidade, que varia entre 0 e 1, necessária para a sua utilização em avaliação económica. Tal não impede, no entanto, de se utilizar os dados dicotómicos obtidos através da técnica DCE conjuntamente com os dados contínuos cTTO, através de modelos híbridos. Quando se considera o uso combinado destas duas técnicas, permite-se a recolha de dados em larga escala através da técnica DCE, sem necessidade da realização de entrevistas presenciais, e a recolha de dados em menor escala de dados cTTO, com recurso a entrevistadores treinados (Rowen et al. 2015). A partir do momento em que se obtém um modelo híbrido consistente e com um bom desempenho, a utilização combinada de dados contínuos e dicotómicos permitirá contornar as desvantagens associadas ao cTTO, relacionadas com a morosidade e dispendiosidade da recolha de dados.

Perante o que foi exposto, conclui-se que nenhuma das técnicas de valoração em análise é perfeita. À luz da investigação actual é possível que não exista, neste momento, nenhuma técnica de valoração perfeita. A escolha por um modelo híbrido baseia-se na assumpção de que os indivíduos possuem uma função única de utilidade. Se as utilidades obtidas através das técnicas cTTO e DCE fossem iguais não haveria necessidade de combinar estas técnicas, excepto para obter coeficientes mais precisos. Dadas as limitações que ambas incorporam, é possível que os “verdadeiros” valores de utilidade não se

consigam obter apenas com recurso a uma das técnicas. Neste caso, qual o impedimento para se recorrer a uma combinação de dados com recurso a modelos híbridos? Com isto, não se pretende defender que as limitações presentes numa técnica são compensadas ou anuladas pelas limitações existentes na outra técnica, obtendo resultados não enviesados. Os estudos efectuados até ao momento não sugerem tal resultado mas apontam na direcção de que as limitações do cTTO não são incrementadas pela DCE, antes pelo contrário. Perante estes argumentos e na ausência actual de medidas estatísticas que possibilitem a avaliação dos modelos híbridos, tal como o erro médio quadrado ou o erro médio ajustado (que não pode ser utilizado por ausência de evidência científica) ou os critérios de informação AIC e BIC (pelo facto da verosimilhança logarítmica de um modelo híbrido ser superior, por definição, à da modelação individual cTTO ou DCE), a decisão pela utilização de um modelo híbrido apenas poderá ser normativa. Estes foram também os argumentos partilhados nos trabalhos apresentados por Nan Luo, Mark Oppe, Juan M. Ramos-Goñi e Kim Rand-Hendriksen na *International Society for Pharmacoeconomics and Outcomes Research 7th Asia-Pacific Conference*, em Setembro de 2016, e por Juan M. Ramos-Goñi no 2nd *EuroQol Academy Meeting*, em Março de 2017.

No entanto, o grupo EuroQol assume que a investigação destes modelos está longe de estar concluída, abrindo múltiplos caminhos de pesquisa tais como a utilização de modelos híbridos destinado a dados em painel, modelos híbridos com coeficientes aleatórios e modelos híbridos derivados com recurso a um número inferior de valorações cTTO.

Com o intuito de melhorar a DCE, vários investigadores que colaboram com o EuroQol, a nível mundial, estão actualmente a dedicar a sua pesquisa ao desenvolvimento do desenho desta técnica e da forma de apresentação das opções de escolha. Uma das linhas de investigação envolve a inclusão de três alternativas de escolha (tríplices), em que duas alternativas respeitam a estados de saúde descritos pelo EQ-5D-5L e a terceira alternativa respeita à morte imediata, a inclusão da duração (tempo) como atributo e uma pergunta de ordenação relativa às alternativas apresentadas (pior/melhor opção).

A figura 43 apresenta um exemplo do que está actualmente a ser testado por investigadores do grupo EuroQol. Uma das áreas de investigação consiste precisamente em

avaliar o impacto da terceira alternativa ser o equivalente a morte imediata, que permite a obtenção de valores melhores e piores que morte, ou correspondente a saúde perfeita.

Figura 43 - Exemplo de investigação futura no desenho da DCE

	Estado 1	Estado 2	Morte Imediata
Mobilidade	Incapaz de andar	Problemas ligeiros em andar	
Cuidados Pessoais	Incapaz de me lavar ou vestir sozinho(a)	Incapaz de me lavar ou vestir sozinho(a)	
Actividades Habituais (por exemplo, trabalho, estudos, actividades domésticas, actividades em família ou de lazer)	Problemas moderados em desempenhar as minhas actividades habituais	Problemas moderados em desempenhar as minhas actividades habituais	
Dor/Mal-estar	Com dores ou mal-estar extremos	Com dores ou mal-estar extremos	
Ansiedade/Depressão	Extremamente ansioso(a) ou deprimido(a)	Ligeiramente ansioso(a) ou deprimido(a)	
Irá viver neste estado de saúde durante este período de tempo e depois morrerá	4 meses	16 meses	
Das três opções, assinale qual a considera ser melhor:	Melhor opção? <input type="radio"/>	Melhor opção? <input type="radio"/>	Melhor opção? <input type="radio"/>
Das três opções, assinale qual a considera ser pior:	Pior opção? <input type="radio"/>	Pior opção? <input type="radio"/>	Pior opção? <input type="radio"/>

Fonte: Exemplo retirado da sessão formativa sobre DCE apresentada na 2nd EuroQol Academy Meeting, em Março de 2017.

Outras correntes de pesquisa consistem na avaliação do impacto do atributo duração e da aplicação de modelos econométricos que permitem incorporar preferências não lineares ao longo do tempo, na imposição de restrições ao desenho da DCE relativas à sobreposição mínima dos níveis de cada atributo e ortogonalidade (independência dos níveis dos atributos), no desenho de opções de escolha com ordem de apresentação aleatória dos atributos e na análise sobre o efeito que distintas formas de apresentação da tarefa de escolha têm na resposta do participante (realce de palavras, cor distinta das opções de resposta).

No fundo, o que se pretende para que a DCE seja considerada uma técnica de valoração por si só é consensualizar um método que produza valores consistentes, com resultados reproduzíveis e comparáveis e que seja aceitável para os utilizadores finais. Para tal suceder, é necessário i) definir os atributos mais adequados a introduzir no conjunto de

escolhas, de modo a permitir a troca entre situações melhores e piores e duração de tempo; ii) consensualizar uma apresentação simples das escolhas no conjunto de opções para resposta; iii) identificar um método estatístico robusto com vista à definição do desenho do conjunto de escolhas e iv) reconhecer uma estrutura a seguir na modelação dos dados que origine utilidades reconhecidas na estrutura dos QALY.

Nesta tese testou-se um conjunto de modelos econométricos, em função do tipo de dados, com o intuito de gerar um sistema de valores para o EQ-5D-5L, representativo das preferências da população portuguesa. Os dados cTTO possibilitaram a obtenção de um sistema de valores a partir do modelo que obteve um melhor desempenho – a regressão por intervalo GLS censurada em 1 (Modelo 5). Pelo que foi exposto anteriormente, os dados obtidos com recurso à técnica de valoração DCE não podem ser utilizados isoladamente para estimar uma tarifa. Não obstante os valores resultantes do modelo de dados cTTO poderem constituir uma resposta para o sistema de valores, pelo facto de se considerar que as técnicas cTTO e DCE incorporam informação distinta mas complementar que poderá melhor representar as preferências da amostra, a opção final recaiu nos resultados estimados por um modelo híbrido heteroscedástico (Modelo 10) que combina estes dois tipos de dados (contínuos e dicotómicos). Considerando os argumentos apresentados anteriormente, a decisão por este modelo teve ainda por base o facto de este estimar coeficientes consistentes e estatisticamente significativos. No encontro académico promovido pelo grupo EuroQol, ocorrido em Março de 2017 na Holanda, foi possível discutir e consensualizar estas opções com os restantes países que terminaram a recolha de dados e se encontravam, tal como o estudo português, na fase de modelação de dados e escolha final do modelo para estimação do sistema de valores, designadamente a Irlanda, Polónia e Malásia.

Para além das questões metodológicas em torno da opção pelo modelo econométrico, há ainda um aspecto relevante relacionado com a exclusão de dados. Nos estudos de valoração de instrumentos de saúde baseados em preferências é comum a aplicação de critérios de exclusão (Engel et al. 2016). Vários critérios podem ser utilizados para excluir a totalidade ou parte dos dados associados a um participante. Attema e colaboradores (2013) identificam os seguintes critérios como possíveis de utilizar com vista

à exclusão de dados: participantes que referiram não ter compreendido a tarefa de valoração, não distinguiram qualquer estado de saúde, utilizaram um número limitado de iterações na valoração de todos os estados de saúde, não trocaram tempo (*non-traders*) e valoraram estados de saúde com problemas ligeiros de forma inferior que estados de saúde severos. No entanto, há que entender que um participante classificado numa destas características não está forçosamente a errar, podendo apenas estar a evidenciar características pessoais ou o efeito do seu contexto familiar e/ou social.

A versão 2.0 do protocolo EQ-VT utilizada neste estudo introduziu o módulo de feedback com o intuito de permitir ao entrevistado rever a ordem das respostas dadas e assinalar algum estado de saúde que eventualmente considere mal posicionado. Este desenvolvimento permitiu considerar os estados assinalados no módulo de feedback como um potencial critério de exclusão.

A forma de gerir os critérios de exclusão é muito variável entres os diversos estudos de valoração publicados, não existindo qualquer norma metodológica consensualizada. A posição formal do grupo EuroQol é que qualquer critério de exclusão pode ser considerado arbitrário (comunicação verbal). As respostas dos participantes que parecem implausíveis aos investigadores podem reflectir um equilíbrio entre a qualidade dos dados e a verdadeira representatividade das preferências societárias, isto é, podem ser perfeitamente consistentes e expectáveis numa amostra representativa da população. Ao invés de reagir a respostas que parecem pouco razoáveis, os investigadores devem garantir, através do treino adequado dos entrevistadores e da explicação cabal de todo o procedimento aos participantes, que este tipo de resposta é minimizado num estudo de valoração e, assim, evitar a aplicação de critérios de exclusão na modelação dos dados.

O estudo de valoração do EQ-5D-3L realizado no Zimbabué, em 2003, incluiu, na modelação dos dados, as respostas dos participantes consideradas como inconsistentes tendo por base que estas nem sempre indicam falta de entendimento dos indivíduos e reconhecendo o direito à inclusão de todos os participantes na medida em que não se pode esperar que todos assumam um comportamento racional nos seus julgamentos sobre a valoração dos estados de saúde (Jelsma et al. 2003). O mesmo sugerem Engel e colaboradores (2016) quando consideram inadequada a exclusão dos indivíduos que não

trocam tempo (*non-traders*) apenas porque isso implica um valor próximo de infinito atribuído à vida, independentemente do estado de saúde que lhes é pedido para valorar. A revisão sistemática da literatura efectuada por estes autores refere que a exclusão deste tipo de indivíduos não é muito comum nos estudos de valoração analisados. Quanto à exclusão de indivíduos que valoram estados de saúde ligeiros de forma inferior relativamente a estados de saúde severos, tal pode meramente indicar incerteza ou o impacto de factores externos do foro pessoal e social e não implicar uma deficiente compreensão da tarefa de valoração (Engel et al. 2016).

Estes motivos, entre outros, levaram Devlin e colaboradores a concluir que não é possível introduzir uma regra de ouro para a aplicação de critérios de exclusão nos estudos de valoração e que qualquer tentativa em fazê-lo deve ser evitada (Devlin et al. 2003). Neste âmbito, a aplicação de qualquer critério de exclusão de dados em estudos de valoração deve ser bem fundamentada e as suas consequências avaliadas através da análise de sensibilidade.

Neste estudo, em 1.451 participantes identificaram-se cinco indivíduos não dispostos a trocar tempo (*non-traders*), cinco indivíduos que valoraram todos os estados de saúde de forma igual (mas distintos de 1) e um indivíduo caracterizado como *funny*. Para além do número de indivíduos com estas características constituir-se como residual (n=11), os argumentos expostos anteriormente levaram a que não se procedesse à sua exclusão neste estudo.

O mesmo não sucedeu com a análise ao módulo de feedback como possível critério de exclusão. Considerando que os 1.119 estados de saúde assinalados pelos participantes, em 14.510 tarefas de valoração, representaram respostas consideradas como mal ordenadas pelos indivíduos, a opção, neste estudo, foi excluir estas valorações. Na medida em que são os próprios indivíduos da amostra que consideram a valoração destes estados de saúde como incorrecta ou mal ordenada, a exclusão dos 1.119 estados de saúde não se afigura como arbitrária. Dado o módulo de feedback ser um desenvolvimento recente do protocolo EQ-VT, introduzido apenas em 2015, são escassos os estudos de valoração realizados com esta componente pelo que, tanto quanto se

conhece, até ao momento actual, nenhum outro estudo de valoração considerou os estados assinalados no módulo de feedback como elegíveis para exclusão.

Importa ainda salientar outros factores que, nos estudos de valoração, podem desempenhar uma influência significativa na qualidade dos dados obtidos. Um dos aspectos concerne ao modo de administração do instrumento de medição em saúde e respectivas técnicas de valoração. As vantagens da aplicação de entrevistas presenciais assistidas por computador ou CAPI, tal como sucedeu neste estudo, sobre questionários aplicados por via postal são reconhecidas. O facto das entrevistas presenciais contarem com a presença de um entrevistador preparado para as realizar, possibilita a preparação e o esclarecimento de dúvidas do participante que de outra forma não seria possível. Este foi também o método seguido nos estudos nacionais de valoração do EQ-5D-3L e SF-6D (Ferreira, Ferreira, Pereira, et al. 2014b; Ferreira, Ferreira, et al. 2010). No entanto, este tipo de entrevistas pode constituir uma limitação significativa em estudos com orçamento limitado. Este constrangimento pode, no entanto, ser contornado se os instrumentos de medição em saúde incluírem no futuro um menor número de tarefas de valoração mais complexas (como o cTTO) ou passarem a utilizar técnicas isoladas mais simples como a DCE. Mesmo que não se preveja, a curto prazo, a utilização individual da DCE, a possibilidade de se poder reduzir o número de tarefas de valoração cTTO quando se combinam estas tarefas de valoração, permite, por si só, reduzir o tempo de entrevista e minimizar custos. Por outro lado, o desenvolvimento de entrevistas guiadas por computador apresenta uma dupla vantagem: em primeiro lugar, normaliza todo o processo de entrevista garantindo que esta é efectuada sempre da mesma forma; em segundo lugar, possibilita a sua implementação *online*, reduzindo custos óbvios com a recolha de dados. Enquanto persistir a dependência em entrevistadores para a obtenção de dados nos estudos de valoração, o seu treino e monitorização constantes configuram uma componente central na qualidade da informação. Não obstante os recursos envolvidos para esta finalidade neste estudo, designadamente o número de horas dedicado a treinos e retreinos de entrevistadores e à produção e análise regulares dos relatórios QC Tool, os resultados aqui obtidos são consequência directa do processo de controlo de qualidade que ocorreu ao longo do período de recolha de dados.

O presente estudo inseriu-se na terceira vaga de estudos de valoração do EQ-5D-5L tendo beneficiado de um conjunto de desenvolvimentos no processo de controlo de qualidade (versão 2.0 do protocolo QC Tool), com vista ao aumento da qualidade dos dados obtidos e que incluiu as tarefas de treino prévias à valoração cTTO, a implementação de janelas de diálogo para confirmação da resposta do participante após cada uma das tarefas de valoração cTTO e DCE e o módulo de feedback. Os melhoramentos utilizados no protocolo EQ-VT foram essenciais para a obtenção de dados válidos com vista à concretização dos objectivos deste estudo. No final, a metodologia seguida permitiu obter um sistema de valores válido para o EQ-5D-5L destinado à população portuguesa em geral.

Não obstante a validade dos resultados obtidos neste estudo há um conjunto de aspectos metodológicos que, se forem contemplados em futuros estudos de valoração, poderão contribuir para a melhoria da qualidade dos dados. Alguns dos aspectos foram já mencionados neste capítulo mas é importante sumariá-los como possíveis limitações deste estudo. Um deles respeita ao desenho da DCE. Este estudo considerou apenas no desenho da DCE os atributos e níveis previstos no sistema descritivo do EQ-5D-5L, não permitindo a utilização independente dos dados de valoração obtidos através desta técnica pelo facto de não utilizar nenhum dos métodos previstos para ancorar os valores obtidos na escala de utilidade do conceito do QALY. Assumindo que, ainda assim, esta técnica de valoração é válida e promissora como técnica de utilização individual, afigura-se um vasto caminho de investigação, por exemplo, no que se refere à introdução do atributo duração (anos, meses, dias), na avaliação da introdução de alternativas como “morte imediata” ou “saúde perfeita” no conjunto de escolhas (tríplices) e na apresentação aleatória dos atributos. Enquanto não for possível utilizar os dados dicotómicos da DCE de forma independente, é igualmente importante aprofundar a investigação de modelos econométricos que integrem dados híbridos e que permitam concluir por estatísticas válidas que permitam a comparação do seu desempenho.

Os resultados obtidos neste estudo constituem as preferências da população portuguesa em geral por estados de saúde descritos pelo instrumento de medição em saúde EQ-5D-5L. No âmbito da decisão nacional sobre a alocação societária de recursos escassos, financiados pelos impostos dos contribuintes, potenciais utilizadores dos

recursos, o sistema de valores obtido neste estudo poderá ser utilizado para informar estudos de avaliação económica em saúde, na medida em que teve por base a auscultação de uma amostra representativa da população portuguesa. No entanto, se se considerar outro nível decisório, como o de avaliação de medicamentos destinados ao tratamento de determinadas patologias (e.g. diabetes), a tarifa de valores agora obtida poderá não representar as preferências dos indivíduos portadores dessas doenças. O mesmo se poderá dizer se se pretender elicitare as preferências de grupos específicos, como crianças ou idosos.

Os sistemas de valores obtidos nos estudos de valoração são, por convenção, baseados na média aritmética como forma de representar as preferências médias da população em geral. No entanto, a base normativa para a selecção desta medida não tem merecido a atenção devida na literatura publicada. Alguns investigadores têm demonstrado algum interesse no estudo de medidas alternativas de tendência central, tal como a mediana, numa perspectiva de mitigar o impacto que os valores extremos associados a alguns indivíduos têm na média de valores (Devlin et al. 2017). No entanto, esta área carece ainda de investigação adicional.

8. CONCLUSÃO

As decisões em saúde são cada vez mais baseadas em estudos de avaliação económica que medem resultados com recurso a QALY. Esta unidade de medida combina, num só indicador, a quantidade de anos de vida de um indivíduo e a qualidade desse período de vida através de pesos baseados em preferências relativas aos estados de saúde sob avaliação. Estes pesos são estimados com recurso a instrumentos de medição em saúde baseados em preferências, tal como o EQ-5D. Os pesos associados a cada estado de saúde constituem o valor relativo que os indivíduos atribuem a viver em cada um desses estados.

Na medida em que as sociedades diferem entre si e possuem valores específicos baseados no seu passado cultural, não basta a determinação de um só sistema de valores para uso partilhado de diversos países. Considerando as características únicas associadas à população de cada país, é necessário a condução de estudos nacionais de valoração, que reflectam as preferências de uma amostra representativa da sua população. Os decisores políticos e as agências de avaliação de tecnologias em saúde de cada país optam pela população em geral como fonte de preferências numa perspectiva de que todos são potenciais utilizadores de recursos em saúde e contribuintes que financiam o Serviço Nacional de Saúde.

São várias as técnicas passíveis de utilizar na valoração de estados de saúde. O EQ-5D-5L incorpora tradicionalmente a técnica cTTO mas, dada a sua complexidade, outra técnica de cariz mais simples – DCE – tem vindo a ser incorporada neste instrumento de medição em saúde como forma de obter dados complementares de valoração. O facto de se tratar de uma técnica de eliciação de preferências com algumas limitações, nomeadamente no que concerne à sua dificuldade em ancorar valores na escala de utilidade necessária na aplicação do conceito do QALY, impede a sua utilização independente em estudos de valoração. A pesquisa futura incide eventualmente na exploração desta técnica, seja para utilização conjunta com outra técnica de valoração, seja para utilização individual, no sentido de apresentar tarefas de valoração cognitivamente

mais simples para os participantes, reduzir a dimensão da amostra necessária nos estudos de valoração e diminuir os custos inerentes.

Este estudo estimou o sistema de valores com base no EQ-5D-5L, para a população portuguesa em geral, tendo por base a utilização conjunta de dados contínuos (cTTO) e dicotómicos (DCE). Os dados cTTO quando utilizados de forma individual permitiram encontrar um bom modelo de explicação atribuído a cada estado de saúde. No entanto, nesta tese parte-se do pressuposto que a aplicação conjunta de uma técnica dicotómica como a DCE pode contribuir para acrescentar mais informação ao modelo. Por esse motivo, optou-se pela utilização do modelo híbrido embora se considere perfeitamente aceitável a aplicação do modelo cTTO. O sistema de valores obtido nesta tese pode ser utilizado no cálculo de QALY, no âmbito da avaliação económica em saúde desenvolvida por organismos nacionais do sector público ou do sector privado da saúde. No entanto, a sua utilização em indivíduos com patologias específicas ou grupos de indivíduos com características particulares (como crianças ou idosos) requer investigação adicional.

A concretização de um estudo desta natureza permite a um economista da saúde providenciar um conjunto exaustivo de informação relativo a valoração em saúde e apresentá-lo aos decisores. No final, a função do economista não deverá ser de optar por uma abordagem ou por outra mas sim apresentar de forma transparente as diversas alternativas existentes, gerar sistemas de valores alternativos, complementados pela informação de estatística sobre o seu desempenho e informar sobre as suas limitações. Ao decisor em saúde cabe a deliberação final sobre quais os valores a utilizar.

REFERÊNCIAS BIBLIOGRÁFICAS

Abellán Perpiñán, J. et al., 2012. Lowering the “floor” of the SF-6D scoring algorithm using a lottery equivalent method. *Health Economics*, 21(11), pp.1271–1285. Disponível em: <http://dx.doi.org/10.1002/hec.1792>.

Andrade, V. et al., 2013. Societal Preferences for EQ-5D Health States from a Brazilian Population Survey. *Value in Health Regional Issues*, 2(2), pp.405–412.

Antunes, P., Ferreira, P. & Ferreira, L., 2017. A Utilização da Experiência de Escolha Discreta na Valoração de Estados de Saúde. *Notas Económicas*, In press.

Attema, A. et al., 2013. Time trade-off: One methodology, different methods. *European Journal of Health Economics*, 14(Supl. 1).

Augustovski, F. et al., 2016. An EQ-5D-5L value set based on Uruguayan population preferences. *Quality of Life Research*, 25(2), pp.323–333.

Augustovski, F. et al., 2013. Lead versus lag-time trade-off variants: Does it make any difference? *European Journal of Health Economics*, 14(Supl. 1).

Baltussen, R. & Niessen, L., 2006. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost effectiveness and resource allocation: C/E*, 4, p.14. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1560167&tool=pmcentrez&rendertype=abstract> [Acedido em 28 de Maio de 2014].

Bansback, N. et al., 2014. Testing a discrete choice experiment including duration to value health states for large descriptive systems: Addressing design and sampling issues. *Social Science and Medicine*, 114, pp.38–48. Disponível em: <http://dx.doi.org/10.1016/j.socscimed.2014.05.026>.

Bansback, N. et al., 2012. Using a discrete choice experiment to estimate health state utility values. *Journal of Health Economics*, 31(1), pp.306–318. Disponível em: <http://dx.doi.org/10.1016/j.jhealeco.2011.11.004>.

Bharmal, M. & Thomas, J., 2006. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. *Value in Health*, 9(4), pp.262–271.

Bliemer, M., Rose, J. & Hess, S., 2008. Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling*, 1(1), pp.98–126. Disponível em: [http://dx.doi.org/10.1016/S1755-5345\(13\)70024-1](http://dx.doi.org/10.1016/S1755-5345(13)70024-1).

Brandeau, M., Sainfort, F. & Pierskalla, W., 2004. *Operations Research and Health Care - A Handbook of Methods and Applications*, Springer US.

Brazier, J. et al., 2004. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ*, 13. Disponível em: <http://dx.doi.org/10.1002/hec.866>.

Brazier, J. et al., 2012. Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. *European Journal of Health Economics*, 13(5), pp.575–587.

Brazier, J. et al., 1998. Deriving a Preference-Based Single Index from the UK SF-36 Health Survey. *Journal of Clinical Epidemiology*, 51(11), pp.1115–1128. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0895435698001036> [Acedido em 14 de Novembro de 2016].

Brazier, J. et al., 2007. *Measuring and Valuing Health Benefits for Economic Evaluation*, 2009th ed. Oxford University Press, ed., Oxford.

Brazier, J., Roberts, J. & Deverill, M., 2002. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21(2), pp.271–292.

Briggs, A., Claxton, K. & Sculpher, M., 2006. *Decision modelling for health economic evaluation*, Oxford: Oxford University Press.

Brooks, R., 2013. *The EuroQol group after 25 years*, Dordrecht ; New York: Springer.

Busschbach, J. et al., 2003. *A comparison of EQ-5D time trade-off values obtained in Germany, United Kingdom, Spain*, R. G. Brooks, F. Rabin, & Charro de, eds. Disponível em: http://dx.doi.org/10.1007/978-94-017-0233-1_9.

Chen, A. et al., 2015. The evolution of the disability-adjusted life year (DALY). *Socio-Economic Planning Sciences*, 49, pp.10–15. Disponível em: <http://dx.doi.org/10.1016/j.seps.2014.12.002>.

Cornes, P., 2012. The economic pressures for biosimilar drug use in cancer medicine. *Targeted oncology*, 7 Supl 1, pp.S57-67. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3291824&tool=pmcentrez&rendertype=abstract> [Acedido em 28 de Junho de 2014].

Craig, B. et al., 2013. US valuation of the SF-6D. *Medical Decision Making*, 33(6), pp.793–803. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3740344&tool=pmcentrez&rendertype=abstract>.

Cunillera, O. et al., 2010. Discriminative capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey. *Quality of Life Research*, 19(6), pp.853–864.

Devlin, N. et al., 2003. Logical inconsistencies in survey respondents' health state valuations - a methodological challenge for estimating social tariffs. *Health Economics*, 12(7), pp.529–544. Available at: <http://doi.wiley.com/10.1002/hec.741> [Accessed April 30, 2017].

Devlin, N. et al., 2016. *Valuing health-related quality of life : An EQ-5D-5L value set for England*. Office of Health Economics, 16(1), pp.1–22.

Devlin, N., Shah, K. & Buckingham, K., 2017. *What is the Normative Basis for Selecting the Measure of "Average" Preferences for Use in Social Choices ?* Office of Health Economics, 17(1), pp.1–21.

Devlin, N. & Sussex, J., 2011. *Incorporating Multiple Criteria in HTA: Methods and Processes*. Office of Health Economics, ed., London.

Dolan, P., 1997. Modeling valuations for EuroQol health states. *Medical care*, 35(11), pp.1095–108. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9366889>.

Drummond, M. et al., 2005. *Methods for the economic evaluation of health care programmes*, Oxford, etc.: Oxford University Press.

Earl, P. & Kemp, S., 1999. *The Elgar companion to consumer research and economic psychology*, Cheltenham; Northampton, MA, USA: Edward Elgar.

Engel, L. et al., 2016. Exclusion Criteria in National Health State Valuation Studies: A Systematic Review. *Medical Decision Making*, 36(7), pp.798–810.

EU & OECD, 2016. *Health at a Glance: Europe 2016*, OECD Publishing. Disponível em: http://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2016_9789264265592-en [Acedido em 14 de Março de 2017].

EuroQol Group, 2005. *EQ-5D concepts and methods: a developmental history*, P. Kind, R. Brooks, & R. Rabin, eds., Dordrecht: Springer.

EuroQol Group, 2007. *EQ-5D value sets: inventory, comparative review, and user guide*, A. Szende, M. Oppe, & N. J. Devlin, eds., Dordrecht: Springer.

EuroQol Group, 1990. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*, 16(3), pp.199–208. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/0168851090904219> [Acedido em 26 de Maio de 2016].

EuroQol Group, 2014. *Self-Reported Population Health: An International Perspective based on EQ-5D*, A. Szende, B. Janssen, & J. Cabases, eds., Dordrecht: Springer Netherlands. Disponível em: <http://link.springer.com/10.1007/978-94-007-7596-1> [Acedido em 2 de Outubro de 2016].

Feng, Y. et al., 2016. New Methods for Modelling EQ-5D-5L Value Sets: An Application to English Data. *Office of Health Economics*, 16(02), pp.1-23.

Ferreira, L., Ferreira, P., et al., 2010. A portuguese value set for the SF-6D. *Value in Health*, 13(5), pp.624–630. Disponível em: <http://dx.doi.org/10.1111/j.1524-4733.2010.00701.x>.

Ferreira, L. et al., 2016. Comparing the performance of the EQ-5D-3L and the EQ-5D-5L in young Portuguese adults. *Health and Quality of Life Outcomes*, 14(1). Disponível em: <http://hqlo.biomedcentral.com/articles/10.1186/s12955-016-0491-x> [Acedido em 29 de Setembro de 2016].

Ferreira, L., Ferreira, P., Pereira, L., et al., 2014a. EQ-5D Portuguese population norms. *Quality of Life Research*, 23(2), pp.425–430.

Ferreira, L. et al., 2013. Exploring the consistency of the SF-6D. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 16(6), pp.1023–31. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24041352> [Acedido em 18 de Junho de 2014].

Ferreira, L., Ferreira, P., Pereira, L., et al., 2014b. The valuation of the EQ-5D in Portugal. *Quality of Life Research*, 23(2), pp.413–423.

Ferreira, L., 2003. Utilidades, QALYs e medição da qualidade de vida. *Revista Portuguesa de Saúde Pública*, 3, pp.51–64.

Ferreira, L., Brito, U. & Ferreira, P., 2010. Qualidade de vida em doentes com asma. *Rev Port Pneum*, (XVI), pp.23–55.

Ferreira, L. & Ferreira, P., 2014. *Health State Values and Country-Specific Value Sets*. In A. C. Michalos, ed. *Encyclopedia of Quality of Life and Well-Being Research*. Dordrecht: Springer Netherlands, pp. 2744–2749. Disponível em: http://dx.doi.org/10.1007/978-94-007-0753-5_3805.

Ferreira, L. & Ferreira, P., 2008. Qualidade de vida em doentes com cataratas. *Oftalmolog*, (32), pp.159–75.

Ferreira, L., Ferreira, P. & Baleiro, R., 2008. Qualidade de vida em doentes com artrite reumatóide. *Acta Reum Port*, (33), pp.341–2.

Ferreira, L., Ferreira, P. & Pereira, L., 2014. Comparing the performance of the SF-6D and the EQ-5D in different patient groups. *Acta Médica Portuguesa*, 27(2), pp.236–45. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24813493>.

Ferreira, L. et al., 2011. Reducing the Floor Effect in the SF-6D: A Feasibility Study. *Applied Research in Quality of Life*, 7(2), pp.193–208. Available at: <http://link.springer.com/10.1007/s11482-011-9149-3> [Accessed June 29, 2014].

Ferreira, P., 2010. *Medição do Estado de Saúde e da Qualidade de Vida*. In 30 anos do Serviço Nacional de Saúde: um percurso comentado. Coimbra: Almedina, pp. 461–484.

Ferreira, P. & Ferreira, L., 2006. A medição de preferências em saúde na população portuguesa. *Revista Portuguesa de Saúde Pública*, (24), pp.5–14.

Ferreira, P., Ferreira, L. & Pereira, L., 2013. Contributos para a Validação da Versão Portuguesa do EQ-5D. *Acta Médica Portuguesa*, 26(6), pp.664–676.

Ferreira, P., Ferreira, L. & Pereira, L., 2015. SF-6D Portuguese population norms. *The European Journal of Health Economics*, 16(3), pp.235–241.

Greene, W., 2003. *Econometric analysis*, 5th ed., Upper Saddle River, N.J: Prentice Hall.

Gujarati, D., 2003. *Basic econometrics*, 4th ed., Boston: McGraw Hill.

Hawthorne, G. et al., 2007. The SF36 Version 2: critical analyses of population weights, scoring algorithms and population norms. *Quality of Life Research*, 16(4), pp.661–673.

Hensher, D., Rose, J. & Greene, W., 2015. *Applied choice analysis*, 2nd edition, Cambridge: Cambridge University Press.

Herdman, M. et al., 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, 20(10), pp.1727–1736.

Hjelmgren, J. & Anell, A., 2007. Population preferences and choice of primary care models: a discrete choice experiment in Sweden. *Health policy (Amsterdam, Netherlands)*, 83(2–3), pp.314–22. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17376559> [Acedido em 31 de Março de 2013].

Horsman, J. et al., 2003. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health and quality of life outcomes*, 1, p.54. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=293474&tool=pmcentrez&rendertype=abstract>.

Huang, I. et al., 2007. US and UK versions of the EQ-5D preference weights: Does choice of preference weights make a difference? *Quality of Life Research*, 16. Disponível em: <http://dx.doi.org/10.1007/s11136-007-9206-4>.

Ikeda, S. et al., 2015. Developing a Japanese version of the EQ-5D-5L value set. *J. Natl. Inst. Public Health*, 64(1), pp.47–55. Disponível em: <http://www.niph.go.jp/journal/data/64-1/201564010008.pdf>.

INFARMED, 1998. *Orientações metodológicas para estudos de avaliação económica de medicamentos*. Infarmed, pp.1–44. Disponível em: http://www.ispor.org/peguidelines/source/orien_metodologicas_eaem.pdf.

Instituto Nacional de Estatística, 2012. *Censos 2011 Resultados Definitivos – Portugal*, Instituto Nacional de Estatística, ed., Lisboa.

Janssen, B. et al., 2013. Introducing the composite time trade-off: A test of feasibility and face validity. *European Journal of Health Economics*, 14(Supl. 1).

Janssen, M. et al., 2013. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research*, 22(7), pp.1717–1727. Disponível em: <http://link.springer.com/10.1007/s11136-012-0322-4> [Acedido em 21 de Maio de 2017].

Jelsma, J. et al., 2003. How do Zimbabweans value health states? *Popul Health Met*, 1. Disponível em: <http://dx.doi.org/10.1186/1478-7954-1-1>.

Johannesson, M., Pliskin, J. & Weinstein, M., 1994. A note on QALYs, time tradeoff, and discounting. *Medical Decision Making*, 14(2), pp.188–193.

Johnson, J. et al., 1998. Valuation of EuroQOL (EQ-5D) health states in an adult US sample. *Pharmacoeconomics*, 13(4), pp.421–433.

Johnson, J. et al., 2005. Valuations of the EQ-5D health states: Are the United States and United Kingdom different? *Med Care*, 43. Disponível em: <http://dx.doi.org/10.1097/00005650-200503000-00004>.

Johnston, J. & DiNardo, J., 2001. *Métodos econométricos*, Lisboa: McGraw-Hill.

Kaplan, R., Bush, J. & Berry, C., 1976. Health status: types of validity and the index of well-being. *Health services research*, 11(4), pp.478–507. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1071947&tool=pmcentrez&rendertype=abstract>.

Kenny, P. et al., 2003. Do participants understand a stated preference health survey? A qualitative approach to assessing validity. *International journal of technology assessment in health care*, 19(4), pp.664–81. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15095772>.

Kim, S. et al., 2016. The EQ-5D-5L valuation study in Korea. *Quality of Life Research*, 25(7), pp.1845–1852. Disponível em: <http://link.springer.com/10.1007/s11136-015-1205-2> [Acedido em 17 de Abril de 2017].

Kjær, T., 2005. A review of the discrete choice experiment - with emphasis on its application in health care. *Health Economics Papers*, (1).

Knies, S. et al., 2009. Utilities of the EQ-5D: Transferable or Not? *PharmacoEconomics*, 27(9), pp.767–779. Disponível em: <http://link.springer.com/10.2165/11314120-000000000-00000> [Acedido em 16 de Outubro de 2016].

Kobelt, G., 2013. *Health Economics : An Introduction to Economic Evaluation*, 2013th ed. OHE, ed., London.

Kontodimopoulos, N. et al., 2009. Comparing SF-6D and EQ-5D utilities across groups differing in health status. *Quality of Life Research*, 18(1), pp.87–97.

Krabbe, P. et al., 2014. Multinational Evidence of the Applicability and Robustness of Discrete Choice Modeling for Deriving EQ-5D-5L Health-State Values. *Medical Care*, 52(11), pp.935–943. Disponível em: [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4196797&tool=pmcentrez&rendertype=abstract%5Cn%3CGo to ISI%3E://WOS:000343930300001](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4196797&tool=pmcentrez&rendertype=abstract%5Cn%3CGo%20to%20ISI%3E://WOS:000343930300001).

Lancsar, E. & Louviere, J., 2008. Conducting Discrete Choice Experiments to Inform Healthcare Decision Making a User's Guide. *Pharmacoeconomics*, 26(8), pp.661–677.

Lancsar, E. & Louviere, J., 2006. Deleting “irrational” responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics*, 15(8), pp.797–811. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16615039> [Accessed March 31, 2013].

Lipscomb, J. et al., 2009. Retaining, and Enhancing, the QALY. *Value in Health*, 12(Supl. 1), pp.18–26.

Martin, A. & Lumley, T., 2000. A Comparison of Standard Trade-Off Scores. *Health Care*, 1, pp.137–147.

Mason, J., 1994. Cost-per-QALY league tables: their role in pharmacoeconomic analysis. *PharmacoEconomics*, 5(6), pp.472–481.

McNeil, B., Weichselbaum, R. & Pauker, S., 1978. Fallacy of the Five-Year Survival in Lung Cancer. *New England Journal of Medicine*, 299(25), pp.1397–1401. Disponível em: <http://www.nejm.org/doi/abs/10.1056/NEJM197812212992506> [Acedido em 13 de Novembro de 2016].

Mehrez, A. & Gafni, A., 1987. An empirical evaluation of two assessment methods for utility measurement for life years. *Socio-Economic Planning Sciences*, 21(6), pp.371–375. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/0038012187900115> [Acedido em 13 de Novembro de 2016].

Mehrez, A. & Gafni, A., 1991. The healthy-years equivalent: how to measure them using the standard gamble approach. *Medical Decision Making*, 11(2), pp.140–146.

Miyamoto, J., 1999. Quality-Adjusted Life Years (QALY) Utility Models under Expected Utility and Rank Dependent Utility Assumptions. *Journal of Mathematical Psychology*, 43(2), pp.201–237. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10366516>.

Monteiro, A. et al., 2014. Design of a discrete choice experiment for preferences elicitation of elderly population towards depression. *Value in Health*, 17(3), p.A197. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1098301514011991> [Acedido em 7 de Novembro de 2016].

Morris, S., Devlin, N. & Parkin, D., 2007. *Economic analysis in health care*, Chichester ; Hoboken, NJ: J. Wiley & Sons.

Mulhern, B. et al., 2016. Valuing Health Using Time Trade-Off and Discrete Choice Experiment Methods: Does Dimension Order Impact on Health State Values? *Value in Health*, 19(2), pp.210–217. Disponível em: <http://dx.doi.org/10.1016/j.jval.2015.11.005>.

Murray, C. & Lopez, A., 1996. *The global burden of disease: a comprehensive assessment of mortality and disability from deceases, injuries and risk factors in 1990 and projected to 2010*. Harvard University Press, 1, pp.1–35.

von Neumann, J. & Morgenstern, O., 1944, 2007. *Theory of games and economic behavior*, Princeton: Princeton University Press. Disponível em: <http://site.ebrary.com/id/10632188> [Acedido em 17 de Abril de 2016].

Nord, E., 1992. An alternative to QALYs: The saved young life equivalent (SAVE). *BMJ*, 305(Outubro 10), pp.875–877.

Nord, E., Daniels, N. & Kamlet, M., 2009. QALYs : Some Challenges. *Value in Health*, 12(Supl. 1), pp.10–15.

Norman, R., Viney, R., et al., 2013. Valuing SF-6D Health States Using a Discrete Choice Experiment. *Medical Decision Making*. 34(6):773-86. Epub 11 de Setembro de 2013.

Norman, R., Cronin, P. & Viney, R., 2013. A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. *Applied Health Economics and Health Policy*, 11(3), pp.287–298.

Oppe, M. et al., 2014. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value in Health*, 17(4), pp.445–453. Disponível em: <http://dx.doi.org/10.1016/j.jval.2014.04.002>.

Oppe, M. et al., 2016. EuroQol Protocols for Time Trade-Off Valuation of Health Outcomes. *Pharmacoeconomics*, 34(10), pp.993–1004. Disponível em: <http://link.springer.com/10.1007/s40273-016-0404-1> [Acedido em 14 de Novembro de 2016].

van Osch, S. et al., 2004. Correcting Biases in Standard Gamble and Time Tradeoff Utilities. *Medical Decision Making*, 24(5), pp.511–517. Disponível em: <http://mdm.sagepub.com/cgi/doi/10.1177/0272989X04268955>.

Patrick, D., Bush, J. & Chen, M., 1973. Methods for Measuring Levels of Well-being for a Health Status Index. *Health Services Research*, 8(3), pp.228–245.

Petrou, S. & Hockley, C., 2005. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Economics*, 14(11), pp.1169–1189. Disponível em: <http://doi.wiley.com/10.1002/hec.1006> [Acedido em 12 de Março de 2017].

Pliskin, J., Shepard, D. & Weinstein, M., 1980. Utility Functions for Life Years and Health Status. *Operations Research*, 28(1), pp.206–224. Disponível em: <http://or.journal.informs.org/cgi/doi/10.1287/opre.28.1.206>.

Rabin, R. et al., 2015. EQ-5D-5L user guide. Basic information on how to use the EQ-5D-5L instrument, (Abril), p.28. Disponível em: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf.

Ramos-Goñi, J. et al., 2016. Introducing a hybrid model that combines continuous and dichotomous responses in a single maximum likelihood function : the hyreg command. *Stata Journal*, Under review.

Ramos-Goñi, J. et al., 2017. Quality Control Process for EQ-5D-5L Valuation Studies. *Value in Health*, 20(3), pp.466–473.

Ramos-Goñi, J. et al., 2014. Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach. *Medical Care*, p.1. Disponível em: <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00005650-900000000-99080> [Acedido em 5 de Outubro de 2016].

Rand-Hendriksen, K. & Augestad, L., 2012. Time trade-off and ranking exercises are sensitive to different dimensions of EQ-5D health states. *Value in Health*, 15(5), pp.777–782. Disponível em: <http://dx.doi.org/10.1016/j.jval.2012.04.002>.

Ratcliffe, J. et al., 2011. Valuing Child Health Utility 9D Health States with a Young Adolescent Sample: A Feasibility Study to Compare Best-Worst Scaling Discrete-Choice Experiment, Standard Gamble and Time Trade-Off Methods. *Applied Health Economics and Health Policy*, 9(1), pp.15–27. Disponível em: <http://link.springer.com/10.2165/11536960-000000000-00000> [Acedido em 7 de Novembro de 2016].

Ravens-Sieberer, U. et al., 2010. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Quality of Life Research*, 19(6), pp.887–897. Disponível em: <http://link.springer.com/10.1007/s11136-010-9649-x> [Acedido em 19 de Abril de 2017].

Robinson, A. & Spencer, A., 2006. Exploring challenges to TTO utilities: Valuing states worse than dead. *Health Economics*, 15(4), pp.393–402.

Robinson, A., Spencer, A. & Moffatt, P., 2015. A Framework for Estimating Health State Utility Values within a Discrete Choice Experiment Modeling Risky Choices. *Medical Decision Making*, 35(3), pp.341–350. Disponível em: <http://mdm.sagepub.com/content/35/3/341.short> <http://www.ncbi.nlm.nih.gov/pubmed/25349189>.

Roseboom, K. et al., 2017. Economic evaluations of health technologies in Dutch healthcare decision-making: a qualitative study of the current and potential use, barriers, and facilitators. *BMC Health Services Research*, 17(1). Disponível em: <http://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-017-1986-9> [Acedido em 14 de Março de 2017].

Rowen, D., Brazier, J. & van Hout, B., 2015. A Comparison of Methods for Converting DCE Values onto the Full Health-Dead QALY Scale. *Medical Decision Making*, 35(3), pp.328–340. Disponível em: <http://mdm.sagepub.com/cgi/doi/10.1177/0272989X14559542> <http://publication/doi/10.1177/0272989X14559542>.

Rulleau, B. & Dachary-Bernard, J., 2012. Preferences, rational choices and economic valuation: Some empirical tests. *The Journal of Socio-Economics*, 41(2), pp.198–206. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1053535711001752> [Acedido em 31 de Março de 2013].

Ryan, M. et al., 2006. Using discrete choice experiments to estimate a preference-based measure of outcome—an application to social care for older people. *Journal of Health Economics*, 25(5), pp.927–944.

Ryan, M., Gerard, K. & Amaya-Amaya, M., 2008. *Using discrete choice experiments to value health and health care*, Dordrecht: Springer.

Sakthong, P., Charoenvisuthiwongs, R. & Shabunthom, R., 2008. A comparison of EQ-5D index scores using the UK, US, and Japan preference weights in a Thai sample with type 2 diabetes. *Health and Quality of Life Outcomes*, 6(1), pp.1–9. Disponível em: <http://dx.doi.org/10.1186/1477-7525-6-71>.

Sculpher, M. et al., 2005. *It's just evaluation for decision-making: recent developments in, and challenges for, cost-effectiveness research*. In *Health policy and economics: opportunities and challenges*. Maidenhead: Open University Press, pp. 8–41.

Seiber, W. et al., 2008. *Quality of Well Being Self-Administered (QWB-SA) Scale User's Manual*. Health Services Research Center: University of California, San Diego.

Shah, K. et al., 2016. An empirical study of two alternative comparators for use in time trade-off studies. *Value in Health*, 19(1), pp.53–59. Disponível em: <http://dx.doi.org/10.1016/j.jval.2015.10.012>.

Shah, K., Tsuchiya, A. & Wailoo, A., 2015. Valuing health at the end of life: A stated preference discrete choice experiment. *Social Science and Medicine*, 124, pp.48–56. Disponível em: <http://dx.doi.org/10.1016/j.socscimed.2014.11.022>.

Sintonen, H., 1994. *The 15-D Measure of Health Related Quality of Life : Reliability, Validity and Sensitivity of its Health State Descriptive System*, National Centre for Health Program Evaluation.

Sintonen, H., 1995. *The 15-D Measure of Health Related Quality of Life. II Feasibility, Reliability and Validity of its Valuation System*, National Centre for Health Program Evaluation.

Stiggelbout, A. et al., 1994. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Medical Decision Making*, 14(1), pp.82–90.

Stolk, E. et al., 2010. Discrete Choice Modeling for the Quantification of Health States : The Case of the EQ-5D. *Value in Health*, 13(8), pp.1005–1013.

Tilling, C. et al., 2010. Protocols for Time Tradeoff Valuations of Health States Worse than Dead: A Literature Review. *Medical Decision Making*, 30(5), pp.610–619. Disponível em: <http://mdm.sagepub.com/cgi/doi/10.1177/0272989X09357475>.

Torrance, G., 1986. Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5(1), pp.1–30.

Torrance, G., Thomas, W. & Sackett, D., 1972. A utility maximization model for evaluation of health care programs. *Health Services Research*, 7(2), pp.118–133.

Towers, I., Spencer, A. & Brazier, J., 2005. Healthy year equivalents versus quality-adjusted life years: the debate continues. *Expert review of pharmacoeconomics & outcomes research*, 5(3), pp.245–54. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19807594>.

Tsevat, J. et al., 1994. Using Health-related Quality-of-life Information: clinical encounters, clinical trials, and health policy. *Journal of General Internal Medicine*, 9(10), pp.576–582.

Tsuchiya, A. et al., 2002. Estimating an EQ-5D population value set: the case of Japan. *Health Econ*, 11. Disponível em: <http://dx.doi.org/10.1002/hec.673>.

Tsuchiya, A. et al., 2014. *Using DCE with duration to examine the robustness of preferences across the five dimensions of the EuroQol instrument: the second paper from the FEDEV project*. EuroQol Group Plenary Proceedings.

Verhoef, L., Haan, A. & van Daal, W., 1994. Risk attitude in gambles with years of life: empirical support for prospect theory. *Medical Decision Making*, 14(2), pp.194–200.

Versteegh, M. et al., 2016. Dutch Tariff for the Five-Level Version of EQ-5D. *Value in Health*, 19(4), pp.343–352. Disponível em: <http://dx.doi.org/10.1016/j.jval.2016.01.003>.

Weinstein, M., Torrance, G., & McGuire, A., 2009. QALYs: The basics. *Value in Health*, 12(Supl. 1), 5–9.

WHO, 1946. *The Preamble of the Constitution of the World Health Organization*, Nova Iorque. Disponível em: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:PREAMBLE+TO+THE+CONSTITUTION+OF+THE+WORLD+HEALTH+ORGANIZATION#0> [Acedido em 14 de Junho de 2014].

Wille, N. et al., 2010. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research*, 19(6), pp.875–886. Disponível em: <http://link.springer.com/10.1007/s11136-010-9648-y> [Accessed April 19, 2017].

Williams, I., Robinson, S. & Dickinson, H., 2012. *Rationing in health care: the theory and practice of priority setting*, Bristol: Policy.

Wittrup-Jensen, K. et al., 2008. Methods, Estimating Danish EQ-5D tariffs using the time trade off (TTO) and visual analog scale (VAS). *Health Economics Papers*, 4, pp.257–292.

Wooldridge, J., 2013. *Introductory econometrics: a modern approach*, 5th ed., Mason, OH: South-Western Cengage Learning.

Xie, F. et al., 2016. A Time Trade-off-derived Value Set of the EQ-5D-5L for Canada. *Medical care*, 54(1), pp.98–105. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4674140&tool=pmcentrez&rendertype=abstract>.

Xie, F. et al., 2014. Eliciting preferences to the EQ-5D-5L health states: Discrete choice experiment or multiprofile case of best-worst scaling? *European Journal of Health Economics*, 15(3), pp.281–288.

Yusof, F., Goh, A. & Azmi, S., 2012. Estimating an EQ-5D value set for Malaysia using time trade-off and visual analogue scale methods. *Value in Health*, 15(1 Supl.), pp.S85–S90. Disponível em: <http://dx.doi.org/10.1016/j.jval.2011.11.024>.

APÊNDICE I

Instruções para os entrevistadores

Estudo de valoração EQ-5D-5L (Versão portuguesa para Portugal)

Versão: 3.2

Data: 30 de junho de 2015

Informações pessoais	3
1. Introdução.....	4
1.1 Boas-vindas e introdução.....	4
1.2 Perguntas de preparação.....	6
2. Exemplo do compromisso em tempo (TTO) da cadeira de rodas: introdução	7
3. Exemplo do compromisso em tempo (TTO) da cadeira de rodas: estados melhores que a morte	8
3.1 Apresentar os elementos que surgem no ecrã	8
3.2 Mostrar o que muda quando se clica no botão “A”	9
3.3 Mostrar o que muda quando se clica no botão “B”	10
3.4 Confirmação de que as instruções foram compreendidas.....	11
4. Exemplo do compromisso em tempo (TTO) da cadeira de rodas: estados piores que a morte	12
4.1 Apresentar a situação de estado “pior que a morte”	12
4.2 Apresentar os elementos no ecrã	14
4.3 Demonstração.....	15
4.4 Informação adicional.....	15
4.5 Estados para treino.....	16
4.6 Confirmação de que as instruções foram compreendidas.....	18
5. Tarefas de compromisso em tempo (TTO) para 10 estados de saúde.....	19
5.1 Primeira tarefa e seguintes	19
5.2 Perguntas de <i>feedback</i>	19
6. Módulo de <i>feedback</i> para confirmar as respostas às tarefas TTO.....	20
6.1 Módulo de <i>feedback</i>	20
6.2 Caixa de comentários	21
7. Tarefas de escolha discreta (DCE) para 7 pares de estados de saúde	22
7.1 Introdução.....	22
7.2 Primeira tarefa e seguintes	23
7.3 Perguntas de <i>feedback</i>	23
8. Conclusão de tarefas	24
8.1 Caixa de comentários	24
8.2 Questionário local	24
Teclas de atalho para tarefas TTO.....	25
Pontos de intervenção	26

Informações pessoais

O estudo de valoração do Grupo EuroQol para o EQ-5D-5L utiliza um programa informático normalizado, designado Tecnologia de Valoração do Grupo EuroQol (EQ-VT). Os inquiridos utilizam este programa informático com a ajuda de um entrevistador. Este documento visa normalizar a ajuda prestada pelo entrevistador disponibilizando um guia escrito.

A interação com os inquiridos é dinâmica: alguns inquiridos compreenderão a tarefa mais depressa do que outros. Tente seguir estas instruções para o entrevistador com o máximo rigor possível.

O entrevistador tem quatro tarefas principais:

1. Explicar a tarefa TTO utilizando o exemplo da cadeira de rodas, que consta do EQ-VT. O entrevistador utiliza este exemplo para ilustrar todos os elementos da tarefa. Os elementos relevantes consistem em:
 - apresentar a tarefa utilizada para os estados considerados melhores que a morte: a tarefa “melhor que a morte”;
 - apresentar a tarefa utilizada para estados considerados piores que a morte: a tarefa “pior que a morte”.
 - explicar a funcionalidade dos botões “A” e “B”;
 - explicar a funcionalidade do botão “A e B são semelhantes”;
 - explicar as ideias subjacentes à pergunta: qual é melhor – Vida A ou Vida B?;
 - explicar os pressupostos subjacentes à tarefa que o inquirido não pode esquecer:
 - o inquirido não pode escolher “acabar com a própria vida”;
 - ao considerar um estado de saúde em particular, não existem tratamentos disponíveis para tornar o estado de saúde mais tolerável (p. ex., analgésicos); o estado de saúde seria sempre vivido exatamente “conforme descrito” na caixa azul;
 - não serão encontrados tratamentos ou cura para a doença durante a vida do inquirido.
2. Certifique-se de que, para cada pergunta, o inquirido lê em voz alta a descrição do estado de saúde EQ-5D-5L. É importante que o inquirido esteja ciente de que sempre que avançar para a pergunta seguinte, será apresentado um estado de saúde diferente.
3. Verifique se o inquirido compreende a tarefa.
4. Garanta que o inquirido tem uma experiência agradável com o estudo ao abordar este tópico difícil – é fundamental ser paciente.

As instruções para os entrevistadores que se seguem focam-se nas instruções verbais a serem dadas aos inquiridos durante a entrevista. Este documento não disponibiliza instruções sobre quaisquer outros aspetos da entrevista – por exemplo, disponibilização de fichas de informações, assinatura de formulários de consentimento, etc. – cujos detalhes apresentam diferenças de país para país, dependendo das exigências dos processos de consentimento de ética de cada país. Cabe ao Investigador Principal a responsabilidade de garantir que todos esses processos são tratados de forma apropriada.

As instruções para o entrevistador são mostradas do lado esquerdo das páginas que se seguem em LETRA MAIÚSCULA. Estas instruções não devem ser lidas ao participante.

As instruções para o participante são mostradas do lado direito em ***“texto a negrito e itálico entre aspas”***. Deve ler estas instruções em voz alta para o participante.

NB: se o inquirido pretender terminar a entrevista em qualquer momento (por qualquer motivo), utilize o botão “terminar sessão” do EQ-VT. Não saia do *browser* sem terminar a sessão. Não esqueça que se terminar uma sessão de entrevista antes de esta chegar ao fim, todos os dados das tarefas de valoração TTO e DCE dessa entrevista perder-se-ão.

1. Introdução

1.1 Boas-vindas e introdução

PARA AS SECÇÕES 1-4 DEVERÁ MANTER O CONTROLO DO COMPUTADOR PORTÁTIL. ISTO SIGNIFICA QUE IRÁ ORIENTAR A DEMONSTRAÇÃO PARA O INQUIRIDO. CERTIFIQUE-SE DE QUE O INQUIRIDO ESTÁ SENTADO E QUE AMBOS CONSEGUEM VER BEM O ECRÃ DO COMPUTADOR PORTÁTIL.

“Nesta entrevista, ser-lhe-á apresentada uma série de perguntas no computador. Vou agora apresentar o estudo em termos genéricos.”

LEIA AO INQUIRIDO A INTRODUÇÃO QUE SURGE NO ECRÃ (MOSTRADA NA FIGURA 1).

Figura 1: Introdução que surge no ecrã

PERSPETIVAS SOBRE A SAÚDE

Obrigado por concordar em participar neste estudo.

Estamos a realizar esta entrevista para permitir melhoramentos no sistema de saúde. Pede-se a elementos da população em geral, na qual está incluído(a), que deem o seu contributo para o estudo.

As perguntas referem-se às diferentes opiniões das pessoas sobre a saúde e a doença. Ser-lhe-á pedido que pense sobre alguns estados de saúde por que pode ou não ter passado na sua vida.

Não há respostas certas ou erradas. Diga apenas o que pensa.

Com base na experiência anterior, o estudo deverá demorar entre 30 e 45 minutos. Caso não se sinta à vontade para responder às perguntas, pode interromper em qualquer momento. As suas respostas serão tratadas com rigorosa confidencialidade.

Está preparado(a) para começar?

← Anterior

Seguinte →

PREPARE O INQUIRIDO PARA AS PERGUNTAS DE PREPARAÇÃO, PARA O EXEMPLO DA CADEIRA DE RODAS E PARA OS ESTADOS PARA TREINO.

“Primeiro, ser-lhe-ão feitas perguntas sobre a sua saúde e algumas perguntas pessoais simples, tais como a sua idade e experiência de doença. De seguida, ser-lhe-á apresentado um exemplo dos tipos de perguntas utilizadas para compreender as suas opiniões sobre diferentes problemas de saúde. Analisaremos este exemplo em conjunto, para que compreenda bem aquilo que terá de fazer. De seguida, terá de responder a três perguntas para treino, para ficar a compreender melhor a tarefa e para mostrar o tipo e as gravidades dos estados de saúde incluídos na experiência.”

DÊ AO INQUIRIDO A POSSIBILIDADE DE FAZER PERGUNTAS E NÃO AVANCE ENQUANTO ELE NÃO ESTIVER PREPARADO.

“Tem dúvidas, ou podemos começar?”

1.2 Perguntas de preparação

ORIENTE O INQUIRIDO PELAS PERGUNTAS DE PREPARAÇÃO:

- CINCO PERGUNTAS SOBRE SAÚDE AUTORREPORTADAS
- ESCALA DE SAÚDE AUTORREPORTADA
- PERGUNTAS PESSOAIS DO INQUIRIDO

NA MAIORIA DOS CASOS, SERÁ MAIS FÁCIL PARA SI MANTER O CONTROLO DO COMPUTADOR PESSOAL NESTA FASE.

SE O INQUIRIDO PERGUNTAR O QUE SIGNIFICA “DOENÇA GRAVE”, DIGA QUE PODE DAR-LHE A INTERPRETAÇÃO QUE BEM ENTENDER.

“Agora vou fazer-lhe algumas perguntas sobre si e sobre a sua saúde.”

2. Exemplo do compromisso em tempo (TTO) da cadeira de rodas: introdução

CERTIFIQUE-SE DE QUE CONTROLA O COMPUTADOR PORTÁTIL PARA AS SECÇÕES 2 A 4.

LEIA AO INQUIRIDO A INTRODUÇÃO AO TTO QUE SURGE NO ECRÃ (MOSTRADA NA FIGURA 2).

Figura 2: Introdução ao TTO que surge no ecrã



PREPARE O INQUIRIDO PARA O EXEMPLO DA CADEIRA DE RODAS.

“Estas perguntas pedem-lhe que compare um estado de perfeita saúde com uma situação de alguns problemas de saúde. Neste exemplo, que abordaremos em conjunto, os problemas de saúde relacionam-se com mobilidade – mais especificamente, a necessidade de uma cadeira de rodas para conseguir deslocar-se.”

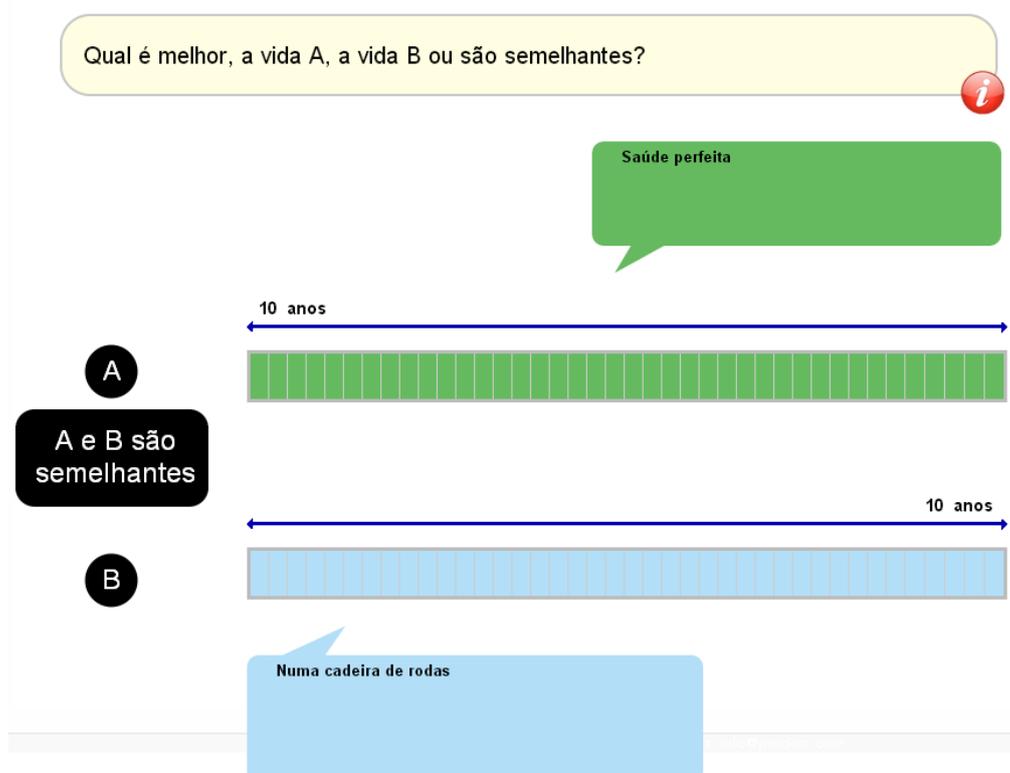
3. Exemplo do compromisso em tempo (TTO) da cadeira de rodas: estados melhores que a morte

3.1 Apresentar os elementos que surgem no ecrã

VERIFIQUE SE OS ELEMENTOS QUE SURGEM NO ECRÃ SÃO CONSISTENTES COM A FIGURA 3.

“Esta é a aparência da tarefa.”

Figura 3: exemplo TTO da cadeira de rodas, Vida A = 10 anos



INDIQUE A BARRA VERDE PARA VIDA A.

“Esta barra verde descreve a designada Vida A. Na Vida A, viverá 10 anos a partir de agora – e durante esses 10 anos, terá uma saúde perfeita. Ao fim dos 10 anos, morrerá.”

SE O INQUIRIDO PERGUNTAR O QUE SIGNIFICA “SAÚDE PERFEITA”, DIGA QUE PODE DAR-LHE A INTERPRETAÇÃO QUE BEM ENTENDER.

INDIQUE A BARRA AZUL PARA VIDA B.

“Esta barra azul descreve a designada Vida B. Na Vida B, também viverá 10 anos, ao fim dos quais morrerá. Porém, nestes 10 anos, terá problemas de mobilidade, o que significa que necessitará de uma cadeira de rodas.”

INDIQUE A PERGUNTA NA CAIXA AMARELA.

Vida B. Qual escolheria? A escolha é entre a Vida A, 10 anos de perfeita saúde, e a Vida B, 10 anos com problemas de mobilidade.”

“Tente imaginar como seria para alguém como você ter de escolher entre a Vida A e a

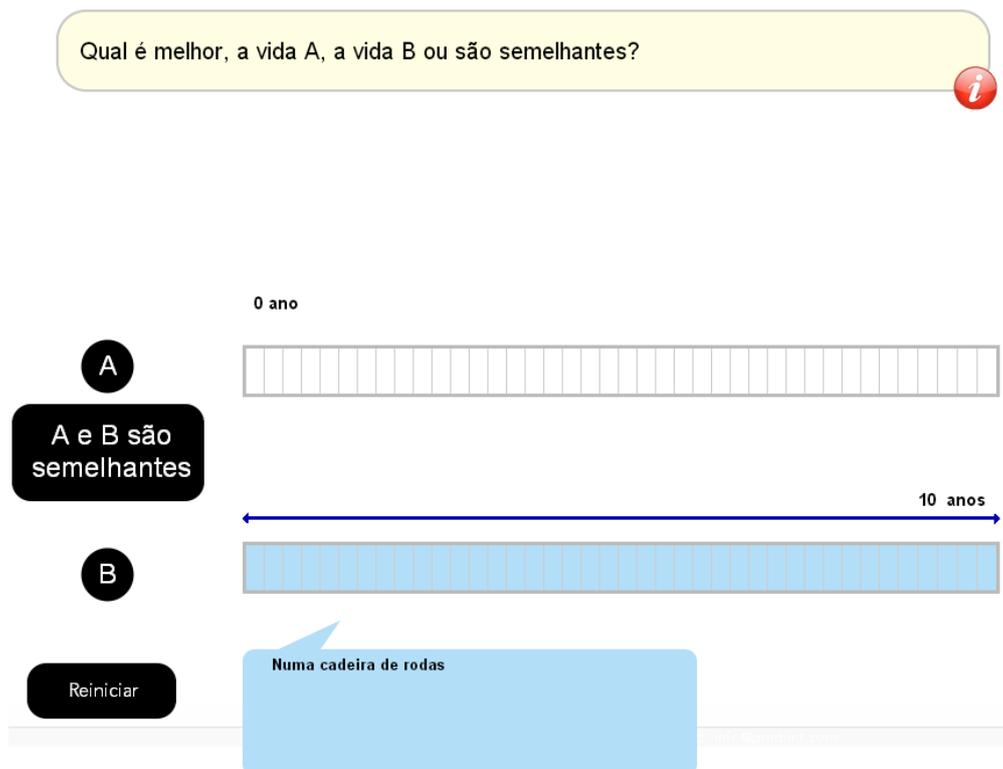
SE O INQUIRIDO RESPONDER QUE PREFERE A VIDA B, OU QUE CONSIDERA QUE A VIDA A E A VIDA B SÃO SEMELHANTES, CONSULTE O PONTO DE INTERVENÇÃO 1 NA PARTE FINAL DESTAS INSTRUÇÕES. CASO CONTRÁRIO, CONTINUE ABAIXO.

3.2 Mostrar o que muda quando se clica no botão “A”

DEMONSTRE COMO ESCOLHER A VIDA A CLICANDO UMA VEZ NO BOTÃO “A”. VERIFIQUE SE OS ELEMENTOS QUE SURGEM NO ECRÃ SÃO CONSISTENTES COM A FIGURA 4.

“Depois de clicar no botão “A”, o computador reage à sua resposta e faz-lhe uma pergunta ligeiramente diferente.”

Figura 4: exemplo TTO da cadeira de rodas, Vida A = 0 anos



INDIQUE A ESCALA VAZIA PARA A VIDA A.

“Ser-lhe-á pedido outra vez que escolha entre a vida A e a Vida B. Desta vez, a escolha é entre zero anos na Vida A (o que significa morrer agora)...”

INDIQUE A ESCALA AZUL PARA VIDA B.

“...ou viver 10 anos a partir de agora numa cadeira de rodas.”

DAR MAIS EXPLICAÇÕES, SE FOR NECESSÁRIO.

SE O INQUIRIDO PARECER CONFUSO QUANTO AO CONCEITO DE VIDA A IMPLICAR “MORRER AGORA”, CONSIDERE A POSSIBILIDADE DE UTILIZAR PALAVRAS ALTERNATIVAS – p. ex.: “PREFERIA MORRER OU VIVER O CENÁRIO APRESENTADO PARA A VIDA B”?

“Se achar que é melhor viver 10 anos numa cadeira de rodas do que morrer agora, deveria escolher a Vida B.”

“Se achar que é melhor morrer agora do que viver 10 anos numa cadeira de rodas (e depois morrer), deveria escolher a Vida A.”

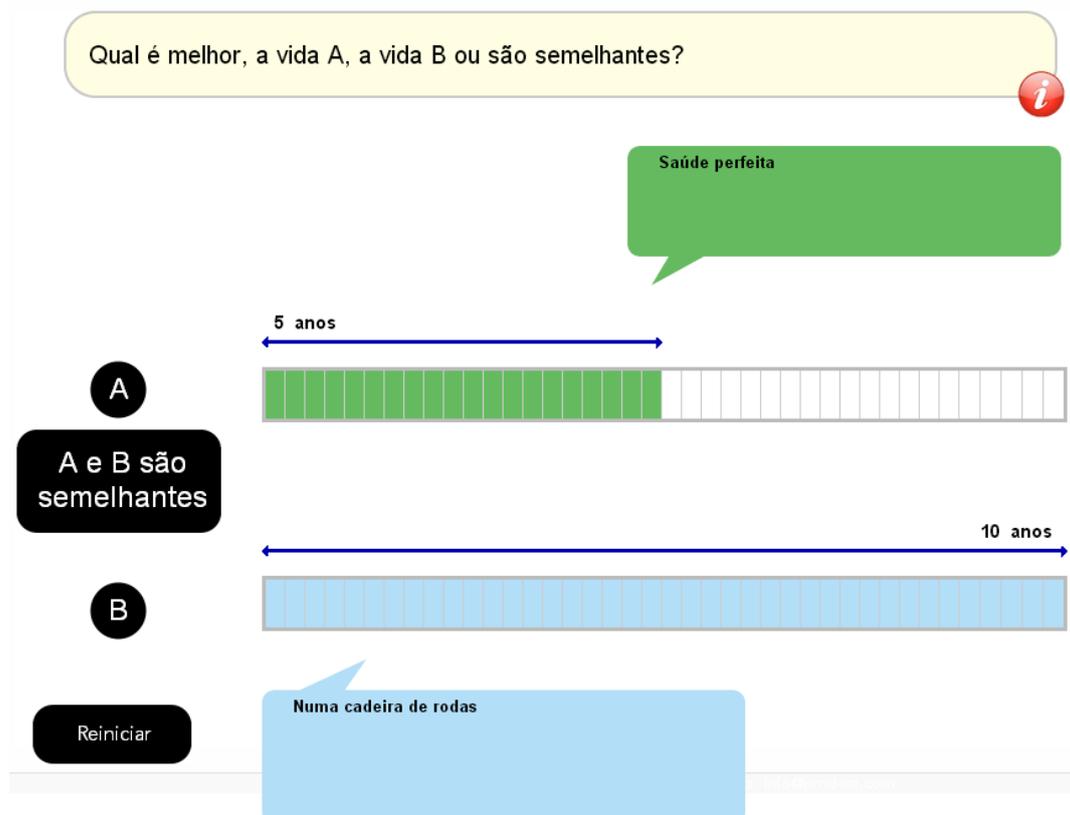
“Que vida escolheria?”

3.3 Mostrar o que muda quando se clica no botão “B”

DEMONSTRE COMO ESCOLHER A VIDA B CLICANDO UMA VEZ NO BOTÃO “B”. VERIFIQUE SE OS ELEMENTOS QUE SURGEM NO ECRÃ SÃO CONSISTENTES COM A FIGURA 5.

“Depois de clicar no botão “B”, o computador reage à sua resposta e faz-lhe uma pergunta ligeiramente diferente.”

Figura 5: exemplo TTO da cadeira de rodas, Vida A = 5 anos



INDIQUE A ESCALA AZUL PARA VIDA B.

“Agora pode escolher entre uma vida mais longa de 10 anos numa cadeira de rodas...”

APONTE PARA O FIM DA ESCALA VERDE PARA VIDA A.

“...e uma vida mais curta de 5 anos num estado de perfeita saúde.”

3.4 Confirmação de que as instruções foram compreendidas

CERTIFIQUE-SE DE QUE O INQUIRIDO COMPREENDE O PROCESSO DE ‘PING PONG’.

SE NECESSÁRIO, DEMONSTRE MELHOR FAZENDO OUTRAS ESCOLHAS E MOSTRANDO COMO O TEMPO DA VIDA A MUDA EM CONSEQUÊNCIA DESTAS ESCOLHAS.

“O que as pessoas escolhem agora depende do que pensam sobre estarem numa cadeira de rodas. Também depende do valor que dão a uma vida longa. Não se esqueça – não há respostas certas ou erradas; estas perguntas apenas pretendem obter a sua opinião.”

“Sempre que faz uma escolha, o tempo da Vida A diminui ou aumenta. Se clicar em “A”, o tempo da Vida A diminui. Se clicar em “B”, o tempo da Vida A aumenta. Estas mudanças no tempo podem ser significativas, podendo chegar aos 5 anos. Porém, à medida que avançar pelas perguntas, as mudanças podem diminuir até 6 meses.”

ATENÇÃO!

Clicar “A e B são semelhantes”. Vai aparecer uma mensagem “A sua resposta sugere que para evitar este estado de saúde durante 10 anos, está disposto(a)”. Responda NÃO para testar a situação de estado de saúde pior que morte.

Clicar “Reiniciar”.

Imaginemos agora que o estado de saúde da pessoa que está em cadeira de rodas é extremamente mau para o considerar pior que morte.

Clique A.

4. Exemplo do compromisso em tempo (TTO) da cadeira de rodas: estados piores que a morte

4.1 Apresentar a situação de estado “pior que a morte”

PREPARE O INQUIRIDO PARA A DEMONSTRAÇÃO DE “PIOR QUE A MORTE”.

“Pelo que vimos até agora, subentende-se que prefere viver 10 anos numa cadeira de rodas do que morrer agora.”

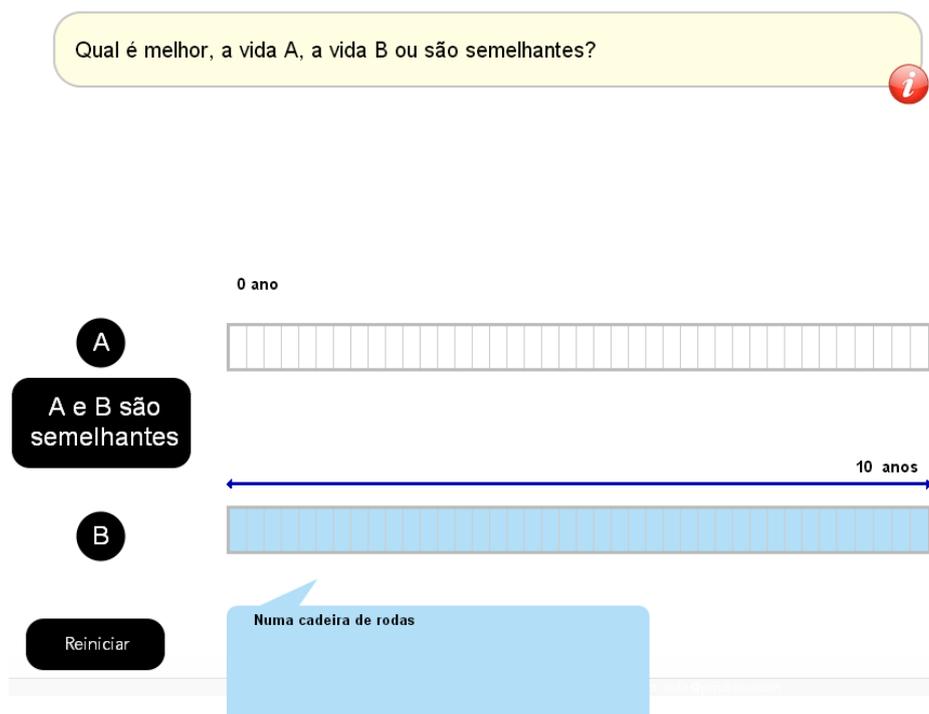
“Ser-lhe-á pedido que considere diversos estados de saúde diferentes. Alguns podem descrever problemas ligeiros, enquanto outros podem descrever problemas de saúde que considera serem realmente maus.”

“Agora, quero mostrar-lhe o que acontece caso se depare com um estado de saúde que é tão mau que preferia morrer agora a ter de viver nesse estado de saúde durante 10 anos.”

DEMONSTRE COMO CHEGAR AO PONTO EM QUE O PERÍODO DE TEMPO NA VIDA A É DE 0 ANOS (CONFORME SE MOSTRA NA FIGURA 6) CLICANDO NO BOTÃO “A”.

“Nesse caso, clicaria no botão “A” porque acha que a Vida “A” é melhor do que viver 10 anos no estado de saúde mau.”

Figura 6: Exemplo do compromisso em tempo da cadeira de rodas, Vida A = 0 anos



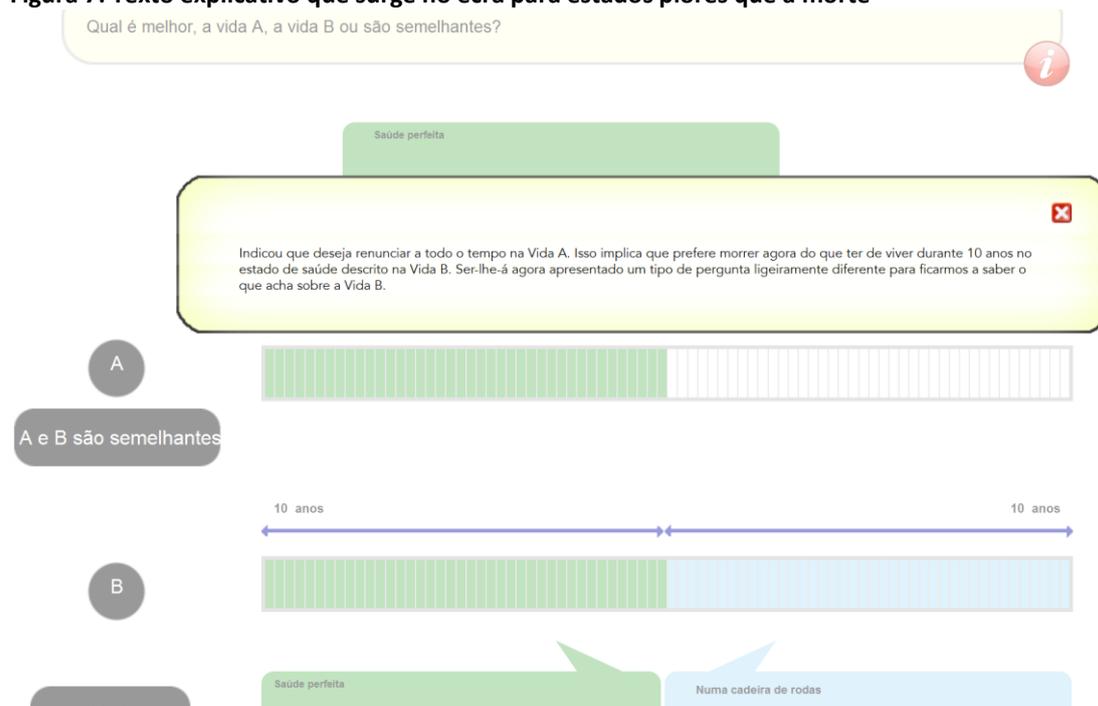
DAR MAIS EXPLICAÇÕES, SE FOR NECESSÁRIO.

SE O INQUIRIDO ESTIVER PRONTO PARA AVANÇAR, CLIQUE UMA VEZ NO BOTÃO “A”. ISTO FAZ APARECER UMA CAIXA DE POP-UP QUE APRESENTA A TAREFA TTO PARA ESTADOS PIORES QUE A MORTE (CONFORME SE MOSTRA NA FIGURA 7).

“Agora, pode escolher entre zero anos na Vida A (o que significa morrer agora) e 10 anos de vida com problemas na Vida B.”

“Se os problemas de saúde em causa forem mesmo maus, pode considerar que preferia morrer do que ter de viver neste estado de saúde durante 10 anos. Nesse caso, deve clicar no botão “A”. ”

Figura 7: Texto explicativo que surge no ecrã para estados piores que a morte



LEIA AO INQUIRIDO O TEXTO EXPLICATIVO QUE SURGE NO ECRÃ PARA ESTADOS PIORES QUE A MORTE (MOSTRADO NA FIGURA 7). FECHER A CAIXA DE POP-UP QUE SURGE NO ECRÃ QUANDO O INQUIRIDO ESTIVER PREPARADO PARA CONTINUAR.

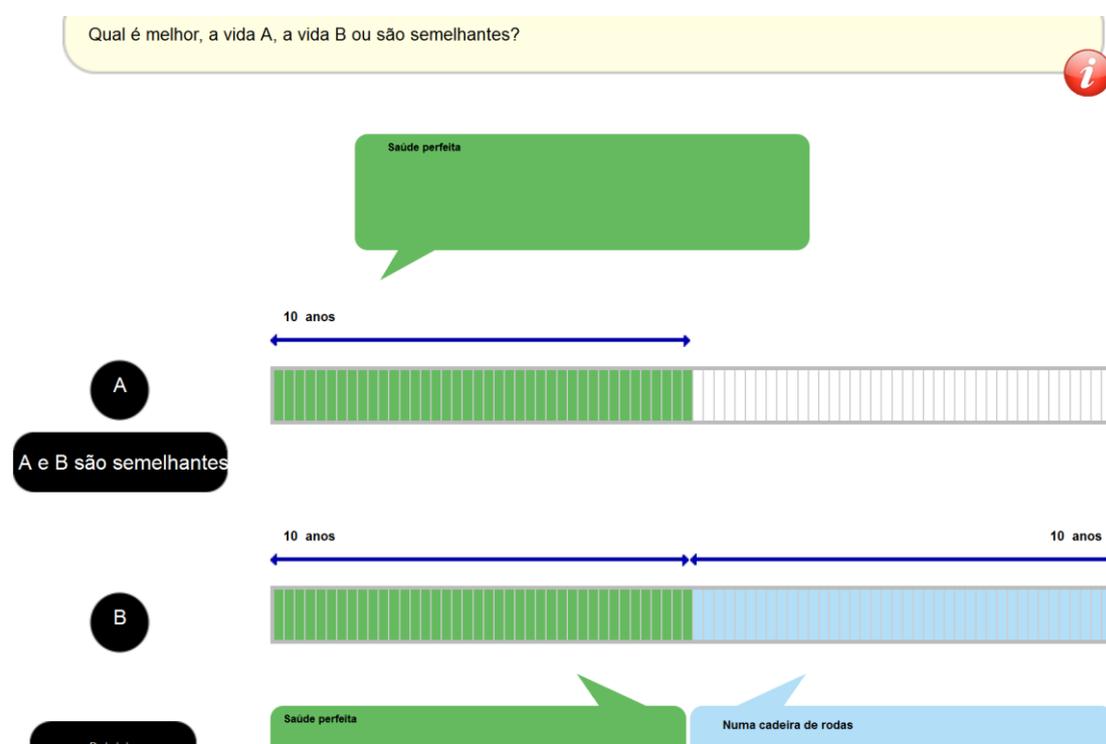
4.2 Apresentar os elementos no ecrã

VERIFIQUE SE OS ELEMENTOS NO ECRÃ SÃO CONSISTENTES COM A FIGURA 8.

EXPLIQUE COMO A TAREFA MUDOU.

“Agora, ser-lhe-á feita uma pergunta ligeiramente diferente. Continua a ter de escolher entre a Vida A e a Vida B, e a barra azul na Vida B continua a referir-se a passar 10 anos com os mesmos problemas de saúde que antes. Contudo, os problemas da Vida B agora não surgem imediatamente, mas sim ao fim de 10 anos num estado de perfeita saúde. Por isso, a Vida B tem agora a duração total de 20 anos a contar de agora: 10 anos num estado de perfeita saúde, seguidos de 10 anos com problemas de saúde.”

Figura 8: Exemplo do TTO da cadeira de rodas, Vida A = 10 anos, estado pior que a morte



INDIQUE O CENTRO DA ESCALA PARA A VIDA A.

“A Vida A também mudou – agora tem a duração de 10 anos. Por isso, agora pode escolher entre 10 anos de vida de perfeita saúde na Vida A...”

INDIQUE A ESCALA PARA A VIDA B.

“...ou 20 anos na Vida B – 10 anos de perfeita saúde seguidos de 10 anos com problemas de saúde.”

4.3 Demonstração

DEMONSTRE O QUE ACONTECE QUANDO O INQUIRIDO ESCOLHE A VIDA A.

“Tal como anteriormente, se achar que a Vida A é melhor, deve clicar no botão “A”. Deste modo, reduzirá o período de tempo num estado de perfeita saúde na Vida A.”

DEMONSTRE O QUE ACONTECE QUANDO O INQUIRIDO ESCOLHE A VIDA B.

“E, tal como anteriormente, se achar que a Vida B é melhor, deve clicar no botão B”. Deste modo, regressará ao tipo de perguntas anterior.”

CONFIRME SE O INQUIRIDO COMPREENDEU, FAZENDO MAIS DEMONSTRAÇÕES, SE NECESSÁRIO.

“O período de tempo na Vida A continuará a mudar sempre que fizer uma escolha, até chegar a uma situação em que a Vida A e a Vida B apresentam níveis de agrado idênticos para si.

“Este ponto será diferente para pessoas diferentes e depende da qualidade do estado de saúde.”

INDIQUE O BOTÃO “A e B são semelhantes”.

“Quando chegar a esse ponto, deve clicar no botão “A e B são semelhantes”. Aparecerá um ecrã de pop-up no qual lhe será pedido que confirme a sua resposta. Será então apresentado um novo estado de saúde.”

4.4 Informação adicional

EXPLIQUE O QUE FAZER CASO O INQUIRIDO COMETA UM ERRO. FAÇA UMA DEMONSTRAÇÃO, SE NECESSÁRIO.

“Caso cometa um erro – por exemplo, se clicar em “A” quando na verdade preferia a Vida B – informe-me. Não é necessário recomeçar do início – pode regressar à posição anterior simplesmente continuando a escolher a vida que acha melhor.

Se clicar por engano no botão “A e B são semelhantes”, poderá regressar à tarefa clicando em “Não” no ecrã de pop-up.

EXPLIQUE AO INQUIRIDO OS PRESSUPOSTOS PRINCIPAIS.

“Antes de começarmos, quero explicar-lhe mais uma coisa.”

“Quando pensa nestas perguntas, só pode escolher entre a Vida A e a Vida B – não existem outras opções. Quando pensa na Vida B, imagine que a Vida B será sempre vivida

exatamente conforme descrito. Não existem tratamentos disponíveis para diminuir os problemas de saúde descritos na caixa azul. A única maneira de evitar esses problemas de saúde é escolher a Vida A. A definição de estado de perfeita saúde na Vida A é aquilo que considerar ser uma saúde perfeita.”

4.5 Estados para treino

4.5.1 Primeiro estado para treino (ligeiro)

PASSE O COMPUTADOR PORTÁTIL AO INQUIRIDO.

APRESENTE AO INQUIRIDO A PRIMEIRA TAREFA TTO.

“Esta pergunta parece semelhante ao exemplo da cadeira de rodas que acabou de ver. Apenas o estado de saúde é diferente.”

INDIQUE O ESTADO DE SAÚDE DESCRITO NA CAIXA AZUL.

“Recorda-se das cinco perguntas sobre a sua saúde a que respondeu no início? O estado de saúde que está a ver agora é uma das maneiras como alguém poderia ter respondido a essas perguntas. Pode ler a descrição?”

CERTIFIQUE-SE DE QUE O INQUIRIDO LÊ ATENTAMENTE A DESCRIÇÃO DO ESTADO DE SAÚDE.

Tal como no exemplo da cadeira de rodas, pergunta-se qual preferiria: 10 anos de perfeita saúde na Vida A seguidos de morte, ou 10 anos no estado de saúde azul seguidos de morte. Faça a sua escolha clicando no botão apropriado.”

ESPERE ATÉ O INQUIRIDO FAZER A SUA ESCOLHA.

“A Vida A implica morrer agora. Qual acha que é melhor, a Vida A ou a Vida B?”

SE O INQUIRIDO PREFERIR A VIDA B, OU CONSIDERAR QUE A VIDA A E A VIDA B SÃO SEMELHANTES, CONSULTE O PONTO DE INTERVENÇÃO 1 NA PARTE FINAL DESTAS INSTRUÇÕES. CASO CONTRÁRIO, CONTINUE ABAIXO.

DEIXE O INQUIRIDO COMPLETAR A PRIMEIRA TAREFA TTO, INTERVINDO APENAS NOS CASOS INDICADOS NA SECÇÃO “Pontos de intervenção” NA PARTE FINAL DESTAS INSTRUÇÕES.

QUANDO O INQUIRIDO CLICAR NO BOTÃO “A e B são semelhantes” E CONFIRMAR A RESPOSTA NO ECRÃ DE POP-UP, COMPROVE QUE

COMPLETOU O PRIMEIRO TREINO E SALIENTE QUE DEVE AGORA CONSIDERAR UM ESTADO DE SAÚDE DIFERENTE.

“Excelente! Concluiu as perguntas do primeiro estado de saúde para treino.”

INDIQUE O NOVO ESTADO DE SAÚDE DESCRITO NA CAIXA AZUL. QUANDO O INQUIRIDO TIVER COMPROVADO QUE ESTÁ PRESTES A INICIAR UMA NOVA TAREFA, DEIXE-O CONFIRMAR A RESPOSTA NA CAIXA DE POP-UP E AVANCE PARA OS SEGUNDO E TERCEIRO ESTADOS PARA TREINO.

4.5.2 Segundo estado para treino (grave)

APRESENTE AO INQUIRIDO A SEGUNDA TAREFA TTO.

CERTIFIQUE-SE DE QUE ELE LÊ ATENTAMENTE A DESCRIÇÃO DO NOVO ESTADO DE SAÚDE.

“O novo estado de saúde pode assemelhar-se de algum modo ao anterior. É descrito utilizando os mesmos cinco aspetos de saúde, mas como pode constatar, este estado de saúde é muito pior do que o anterior.”

“Cada estado de saúde que vir será diferente. Certifique-se de que lê atentamente as descrições e tente imaginar como seria para alguém como você viver nesse estado de saúde. Sempre que for apresentado um novo estado de saúde, vou-lhe pedir que leia em voz alta a descrição para que eu possa confirmar que compreendeu o novo estado de saúde. Pode ler em voz alta agora?”

DEIXE O INQUIRIDO COMPLETAR A SEGUNDA TAREFA TTO; INTERVENHA NOS CASOS ASSINALADOS NA SECÇÃO “Pontos de intervenção” NA PARTE FINAL DESTAS INSTRUÇÕES.

4.5.3 Terceiro estado para treino (intermédio)

PARA TODAS AS TAREFAS TTO, NÃO SE ESQUEÇA DE:

- CHAMAR A ATENÇÃO DO INQUIRIDO PARA O FACTO DE O ESTADO DE SAÚDE TER MUDADO
- PEDIR AO INQUIRIDO QUE LEIA EM VOZ ALTA A NOVA DESCRIÇÃO
- INCENTIVAR O INQUIRIDO A “PENSAR EM VOZ ALTA”, CASO ELE CONSIDERE QUE ISSO É ÚTIL

Este é o último estado para treino. Leia em voz alta.

Consegue imaginar a vida neste estado? Este estado de saúde pode ser bastante difícil de imaginar. Contudo, isso não quer dizer que o estado de saúde não possa existir”. Por exemplo, neste caso, pode descrever uma pessoa que ficou com os dois braços inutilizados. Isso impediria a pessoa de se lavar e vestir sozinha, bem como de realizar as atividades normais, mas não tem consequências em termos de mobilidade, dor ou nível de ansiedade/depressão.

4.6 Confirmação de que as instruções foram compreendidas

VERIFIQUE SE O INQUIRIDO COMPREENDEU AS INSTRUÇÕES E DÊ MAIS EXPLICAÇÕES, SE NECESSÁRIO.

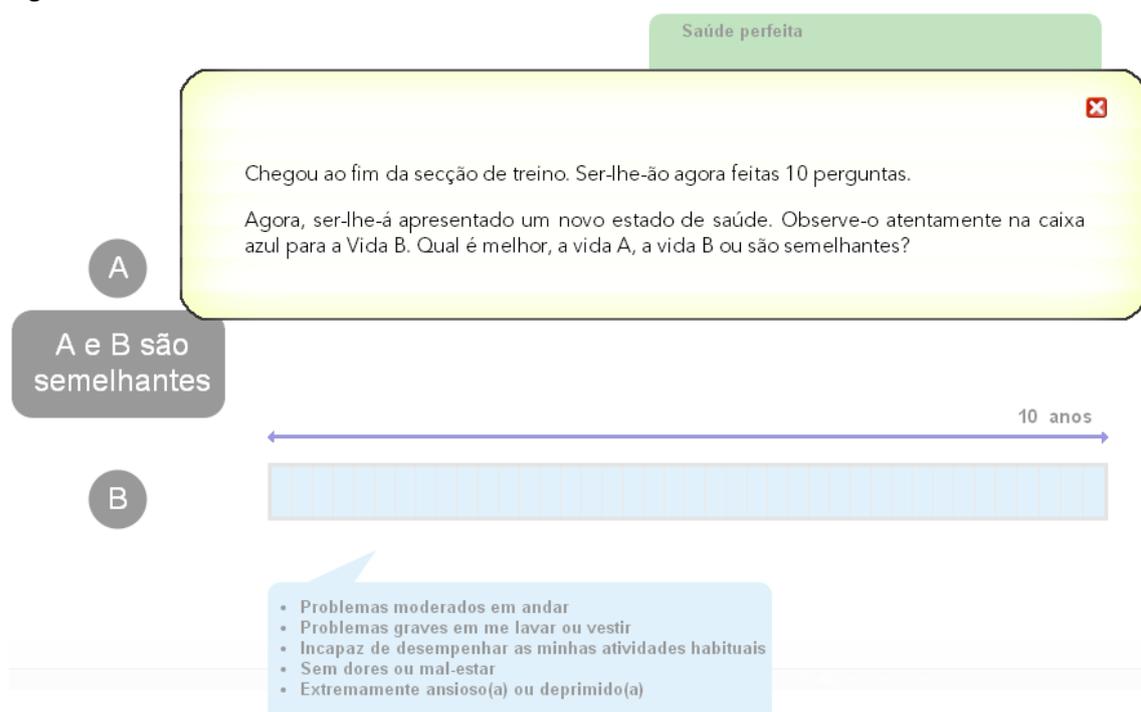
PASSE O COMPUTADOR PORTÁTIL AO INQUIRIDO. SE ACHAR QUE SERIA VANTAJOSO PARA ELE PRATICAR A UTILIZAÇÃO DO COMPUTADOR PORTÁTIL E FAMILIARIZAR-SE COM A TAREFA TTO, PERMITA QUE O FAÇA.

CASO CONTRÁRIO, PEÇA-LHE PARA CLICAR NO BOTÃO “A e B são semelhantes”, O QUE ATIVARÁ UMA CAIXA DE POP-UP QUE APRESENTA AO INQUIRIDO O PRIMEIRO ESTADO DE SAÚDE A SER AVALIADO (CONFORME SE MOSTRA NA FIGURA 9).

“Ser-lhe-á pedido que responda a perguntas sobre 10 estados de saúde diferentes. Considere atentamente cada estado de saúde e pondere bem as suas respostas. Sempre que aparecer um novo estado de saúde, pedir-lhe-ei que leia a descrição do estado de saúde antes de começar a responder às perguntas.”

“Não se esqueça, não há respostas certas ou erradas – apenas pretendemos obter a sua opinião.”

Figura 9: Primeira tarefa TTO



5. Tarefas de compromisso em tempo (TTO) para 10 estados de saúde

5.1 Primeira tarefa e seguintes

PARA TODAS AS TAREFAS TTO, NÃO SE ESQUEÇA DE:

- CHAMAR A ATENÇÃO DO INQUIRIDO PARA O FACTO DE O ESTADO DE SAÚDE TER MUDADO
- PEDIR AO INQUIRIDO QUE LEIA EM VOZ ALTA A NOVA DESCRIÇÃO
- INCENTIVAR O INQUIRIDO A “PENSAR EM VOZ ALTA”, CASO ELE CONSIDERE QUE ISSO É ÚTIL

5.2 Perguntas de *feedback*

QUANDO TERMINAR A TAREFA TTO FINAL (DÉCIMA), DEIXE O INQUIRIDO COMPLETAR AS TRÊS PERGUNTAS DE *FEEDBACK* TTO.

“Gostaríamos de obter o seu feedback sobre as perguntas a que acabou de responder.”

6. Módulo de *feedback* para confirmar as respostas às tarefas TTO

6.1 Módulo de *feedback*

EXPLIQUE AO INQUIRIDO OS ELEMENTOS DO ECRÃ. A FIGURA 10 MOSTRA UM EXEMPLO.

“Aqui pode ver os 10 estados de saúde que acabou de avaliar. Os estados de saúde estão ordenados com base no número de anos na Vida A quando clicou no botão “A e B são semelhantes”.”

“Os estados de saúde que considerou serem os melhores – abdicando de um reduzido número de anos – são mostrados na parte de cima do ecrã. Os estados de saúde que considerou serem piores – abdicando de um elevado número de anos – são indicados na parte de baixo do ecrã. Os estados de saúde a que foi atribuído o mesmo valor são apresentados lado a lado.”

“As caixas azuis mais escuras são utilizadas nos casos em que indicou que é melhor morrer agora do que ter de viver nesse estado de saúde durante 10 anos.”

Figura 10: Exemplo de módulo de *feedback*



CONVIDE O INQUIRIDO A CONSIDERAR SE OS ESTADOS DE SAÚDE ESTÃO ORDENADOS CORRETAMENTE.

INCENTIVE O INQUIRIDO A “PENSAR EM VOZ ALTA” E A DESCREVER OS MOTIVOS POR QUE DESEJA RECONSIDERAR AS SUAS RESPOSTAS.

CASO IDENTIFIQUE ALGUMA DAS SEGUINTE SITUAÇÕES NA ORDENAÇÃO:

- PIOR ESTADO DE SAÚDE COM UMA ORDENAÇÃO MAIS ELEVADA (i.e. VALORADO COMO MELHOR) DO QUE OUTROS ESTADOS DE SAÚDE;
- ESTADOS DE SAÚDE GRAVES COM UMA ORDENAÇÃO MAIS ELEVADA (i.e. VALORADOS COMO MELHORES) DO QUE ESTADOS DE SAÚDE LIGEIOS;
- ELEVADO NÚMERO DE ESTADOS DE SAÚDE COM A MESMA ORDENAÇÃO;

ENTÃO, DEVERÁ REVER TODA A ORDENAÇÃO EM CONJUNTO COM O INQUIRIDO. ISTO IMPLICA DISCUTIR A POSIÇÃO DE CADA ESTADO DE SAÚDE RELATIVAMENTE AOS QUE SE ENCONTRAM EM POSIÇÕES ADJACENTES. PEÇA AO INQUIRIDO QUE CONFIRME SE OS ESTADOS DE SAÚDE EM QUESTÃO SÃO EFETIVAMENTE MELHORES, PIORES OU IDÊNTICOS AOS ESTADOS DE SAÚDE A QUE FOI ATRIBUÍDA UMA ORDENAÇÃO MAIS BAIXA, MAIS ELEVADA OU IDÊNTICA.

“Desejamos saber se está satisfeito com a nossa interpretação das suas respostas. Considere atentamente a ordenação. Depois de refletir, gostaria de reconsiderar a sua opinião sobre algum dos estados de saúde? Se assim for, clique nas caixas de estado de saúde relevantes.”

6.2 Caixa de comentários

QUANDO O INQUIRIDO CONCLUIR O MÓDULO DE *FEEDBACK*, SERÁ CONVIDADO A DEIXAR UM COMENTÁRIO NA CAIXA DE TEXTO. NÃO É OBRIGATÓRIO. O INQUIRIDO PODE CONSULTAR A ORDENAÇÃO CLICANDO NO BOTÃO “MOSTRAR ORDENAÇÃO”, MAS NÃO PODERÁ SELECIONAR NEM ANULAR A SELEÇÃO DE ESTADOS DE SAÚDE NESTA FASE.

7. Tarefas de escolha discreta (DCE) para 7 pares de estados de saúde

7.1 Introdução

VERIFIQUE SE OS ELEMENTOS NO ECRÃ SÃO CONSISTENTES COM A FIGURA 11. LEIA AO INQUIRIDO A INTRODUÇÃO À DCE QUE SURGE NO ECRÃ.

“De seguida, ser-lhe-ão colocadas sete perguntas de um tipo diferente.”

Figura 11: Introdução à DCE que surge no ecrã

Qua

Em cada uma das perguntas que se seguem serão apresentados dois estados de saúde diferentes. Observe-os atentamente. Escolha aquele que, na sua opinião, é melhor premindo o botão “A” ou “B”.

- Problemas ligeiros em andar
- Sem problemas em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar extremos
- Gravemente ansioso(a) ou deprimido(a)

A

- Problemas graves em andar
- Sem problemas em me lavar ou vestir
- Problemas moderados em desempenhar as minhas atividades habituais
- Dores ou mal-estar ligeiros
- Não estou ansioso(a) ou deprimido(a)

B

APRESENTE A TAREFA DCE.

“As perguntas são bastante diferentes das que viu até agora. Mais uma vez, é-lhe pedido que escolha entre A e B. Porém, nestas perguntas, A e B não representam vidas com determinada duração, referem-se simplesmente a dois estados de saúde diferentes. Tal como anteriormente, os estados de saúde são descritos utilizando cinco aspetos de saúde diferentes.”

REALCE A IMPORTÂNCIA DE LER ATENTAMENTE AS DESCRIÇÕES DOS DOIS ESTADOS DE SAÚDE.

“Cada par de estados de saúde apresentado será diferente. Certifique-se de que lê atentamente as duas descrições, e tente imaginar como seria para alguém como você viver nesses estados de saúde. Para cada par de estados de saúde, pretendemos saber qual deles acha melhor.”

7.2 Primeira tarefa e seguintes

PEÇA AO INQUIRIDO QUE LEIA EM VOZ ALTA AS DESCRIÇÕES DOS DOIS ESTADOS DE SAÚDE.

“Apresentamos aqui dois estados de saúde diferentes. Leia em voz alta as duas descrições.”

“Que estado de saúde considera ser melhor, A ou B? Indique a sua resposta clicando no botão preto relevante.”

DEIXE O INQUIRIDO FAZER A SUA ESCOLHA.

DE CADA VEZ QUE FOR APRESENTADO UM NOVO PAR DE ESTADOS DE SAÚDE, PEÇA AO INQUIRIDO QUE LEIA EM VOZ ALTA AS DESCRIÇÕES, DEPOIS DEIXE-O COMPLETAR A TAREFA SOZINHO, ATÉ CONCLUIR A TAREFA DCE FINAL.

7.3 Perguntas de *feedback*

QUANDO TERMINAR A TAREFA DCE FINAL (SÉTIMA), DEIXE O INQUIRIDO COMPLETAR AS TRÊS PERGUNTAS DE *FEEDBACK* DCE.

“Mais uma vez, gostaríamos de obter o seu feedback sobre as perguntas a que acabou de responder.”

8. Conclusão de tarefas

8.1 Caixa de comentários

PEÇA AO INQUIRIDO PARA DEIXAR UM COMENTÁRIO NA CAIXA DE TEXTO, CASO ASSIM O DESEJE. NÃO É OBRIGATÓRIO.

8.2 Questionário local

QUANDO O QUESTIONÁRIO PRINCIPAL ESTIVER CONCLUÍDO, NÃO SE ESQUEÇA DE CARREGAR O QUESTIONÁRIO LOCAL ADICIONAL PARA RECOLHER INFORMAÇÕES SOBRE AS CARACTERÍSTICAS PESSOAIS DO INQUIRIDO.

Teclas de atalho para tarefas TTO

É possível completar as tarefas TTO e DCE utilizando apenas o rato. Contudo, se o inquirido tiver dificuldade em utilizar o rato, estão disponíveis as seguintes teclas de atalho:

Ação	Tecla
Escolher: "A"	a
Escolher: "B"	b
Escolher: "A e B são semelhantes" (apenas TTO)	s
Fechar janela do texto de ajuda	x
Repor pergunta (apenas TTO)	r

Pontos de intervenção

Quando possível, deixe o inquirido realizar as tarefas sozinho. Pode ser necessário intervir em resposta aos seguintes pontos de intervenção nas tarefas TTO:

1. O inquirido não escolhe Vida A quando Vida A envolve 10 anos de perfeita saúde

Pergunte ao inquirido porque fez esta escolha:

- Foi por falta de compreensão?
- Estava a tentar terminar a tarefa rapidamente?

Se necessário, esclarecer – por exemplo, explicar que a sua escolha implica que considera que é melhor estar doente do que de saúde perfeita, ou que estar doente durante 10 anos e estar de saúde perfeita durante 10 anos é a mesma coisa. Peça ao inquirido para continuar apenas quando tiver a certeza de que ele compreendeu totalmente o que é necessário.

2. O inquirido apenas escolhe de modo consistente anos de vida inteiros

Lembre ao inquirido que é possível escolher valores intermédios (por exemplo, é possível escolher 9 anos e 6 meses, etc.).

3. O inquirido tem dificuldade em alternar entre formatos TTO para estados melhores / piores que a morte

De cada vez que o questionário alternar para o formato “estados piores que a morte”, certifique-se de que o inquirido compreende como o formato mudou. Se necessário, dê mais explicações, utilizando a secção 4 como guia.

- Se for evidente que o inquirido de facto não considera que o estado de saúde é pior que a morte, ao clicar no botão “B” fará com que regresse ao formato TTO padrão para estados melhores que a morte.
- Se o inquirido alternar continuamente entre os dois formatos (i.e., porque está indeciso entre preferir a Vida A nos estados melhores que a morte e preferir a Vida B nos estados piores que a morte), sugira-lhe que clique no botão “A e B são semelhantes” para terminar a tarefa, ou peça-lhe que esclareça se acha que o estado é melhor ou pior que zero, e, consoante a resposta, sugira o formato que ele deve utilizar.

4. O inquirido não consegue chegar ao ponto em que considera que a Vida A e a Vida B são semelhantes

Pode dar-se o caso de o ponto de indiferença não ser abrangido pelo *software* (o questionário não permite movimentos inferiores a seis meses). Nesse caso, explique que deve clicar no botão “A e B são semelhantes” para avançar para a tarefa seguinte. Certifique-se de que o inquirido não fica numa situação interminável de escolha, incentivando-o a clicar no botão “A e B são semelhantes” para a escolha que está mais próxima do ponto em que não consegue escolher entre a Vida A e a Vida B.

5. O inquirido clica no botão “A e B são semelhantes” quando a Vida A envolve 0 anos (geralmente na segunda fase de determinada tarefa TTO)

Verifique se o inquirido genuinamente considera que morrer agora (Vida A = 0 anos) é igualmente desejável a viver durante 10 anos no estado de saúde (Vida B). Se o inquirido preferir “morrer” a ter de viver durante 10 anos no estado de saúde, deve clicar no botão “A” e utilizar o formato “estados piores que a morte”. Aparecerá um ecrã de pop-up para verificar a resposta do inquirido.