Avaliação de estados de saúde recorrendo à abordagem MACBETH e comparação com o método Time Trade-Off

Andreia Cristina Sanganha Agostinho

Dissertação para obtenção do Grau de Mestre em

Engenharia Biomédica

Orientadores: Prof. Doutora Mónica Duarte Correia de Oliveira

Dr. Paulo Jorge de Morais Zamith Nicola

Júri

Presidente: Prof. Doutora Cláudia Alexandra Martins Lobato da Silva

Orientador: Prof. Doutora Mónica Duarte Correia de Oliveira

Vogal: Prof. Doutora Teresa Sofia Sardinha Cardoso

Junho de 2015

Agradecimentos

À Professora Mónica Oliveira e ao Dr. Paulo Nicola, pela oportunidade de realização do presente trabalho e pela orientação a apoio prestado.
À Professora Lara Ferreira, pelo interesse e disponibilidade demonstrados e pela partilha de conhecimentos que ajudaram na realização desta tese.
A todos aqueles que participaram no questionário desenvolvido e que se prontificaram para colaborar na sua distribuição.
Aos meus amigos mais próximos e ao António, pela paciência e pelos momentos de descontração.
Aos meus pais, pela oportunidade e pela motivação constante, e ao resto da minha família, por todo o apoio que sempre me deram.

Resumo

Os anos de vida ajustados pela qualidade (QALYs), uma medida comumente utilizada na avaliação de tecnologias em saúde, incorporam num índice sumário alterações na quantidade e na qualidade de vida. A utilização de QALYs requer a modelação da qualidade de vida dos indivíduos, para diferentes estados de saúde, em pontuações baseadas em utilidades, recorrendo para tal a métodos de elicitação de preferências como o *Visual Analogue Scale*, o *Time Trade-Off* (TTO) e o *Standard Gamble*. Dado que estes métodos numéricos de elicitação estão, geralmente, associados a um elevado esforço cognitivo e os seus resultados podem ser diferenciados de acordo com o grau de numeracia dos indivíduos e da população, neste estudo explorou-se o recurso à abordagem não numérica MACBETH (*Measuring Attractiveness by a Categorical Based Evaluation Technique*) para avaliação de estados de saúde.

Um novo protocolo de elicitação de preferências baseado na abordagem MACBETH (que requer apenas julgamentos qualitativos) foi desenvolvido e testado através de um questionário *online,* simultaneamente com o método TTO, numa amostra da população geral portuguesa (n=243). O questionário *online* pediu aos indivíduos para avaliar diretamente 25 estados diferentes, descritos de acordo com o EQ-5D, recorrendo aos dois métodos, bem como, para identificar o método a que preferiram responder. A numeracia dos indivíduos também foi avaliada.

Como resultados principais, determinou-se que: os valores médios derivados do MACBETH e do TTO estão fortemente correlacionados (*r=0.962)*; indivíduos com maiores e menores níveis de numeracia preferiram expressar julgamentos de valor com o MACBETH; e os valores médios obtidos diferem com base no nível de numeracia da população.

Os resultados obtidos sugerem, assim, que vale a pena considerar o recurso a métodos não numéricos de elicitação de preferências em saúde. Pesquisas futuras devem ser direcionadas para a compreensão de eventuais vantagens de métodos não numéricos em populações com doenças específicas e entre populações de diferentes países.

Palavras-chave: QALYs; avaliação de estados de saúde; instrumentos baseados em preferências; MACBETH; TTO.

Abstract

Quality-adjusted life years (QALYs), a measure commonly used in health technology appraisals, incorporates in a summary index changes in quantity and in quality of life. The use of QALYs requires the modelling of individuals' quality of life for distinct health states into utility-based QALY scores, with preference elicitation methods such as the Visual Analogue Scale, the Time Trade-Off (TTO) and the Standard Gamble. Given that these numerical elicitation methods may be associated with a high cognitive effort and their results may be differentiated by individual's and population's numeracy, in this study we explore the use of the MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) non-numerical approach to evaluate health states.

A new protocol for preference elicitation based on the MACBETH approach (only requiring qualitative judgments) was developed and tested in a web survey alongside with TTO in a sample of the Portuguese general population (n=243). The web-survey asked individuals to value directly 25 different EQ-5D health states with the two methods, as well as for their preferred method. Individuals' numeracy was also assessed.

As key results, we found out that: the mean values derived from MACBETH and TTO are strongly correlated (Pearson r = 0.962); participants with higher and lower levels of numeracy preferred expressing value judgments with MACBETH; and mean values obtained differed on the basis of numeracy level.

Results suggest that it is worth considering the use of non-numerical preference elicitation methods in health. Further research may be directed to understand eventual advantages of non-numerical methods in special disease populations, diseases states and across population from different countries.

Keywords: QALYs; preference-based instruments; health states valuation; MACBETH; TTO

Índice

1. Introdução ... 1
1.1. Motivação ... 1
1.2. Objetivos ... 2
1.3. Estrutura do Documento ... 3
2. Contexto ... 4
2.1. Avaliação de Tecnologias de Saúde ... 4
2.2. Avaliação Económica em Saúde ... 6
2.2.1. Análise Custo-Benefício ... 8
2.2.2. Análise Custo-Efetividade ... 9
2.2.3. Análise Custo-Utilidade ... 10
2.3. Medição de Ganhos em Saúde através do uso de QALYs ... 11
2.3.1. Conceito de QALY ... 11
2.3.2. Cálculo dos QALYs ... 12
2.3.3. Limitações ... 12
2.4. Conceito de Utilidade, Valor e Preferência ... 13
2.5. Medição da Qualidade de Vida Relacionada com a Saúde ... 14
3. Revisão Bibliográfica ... 16
3.1. Instrumentos Diretos de Obtenção de Utilidades ... 16
3.1.1. *Visual Analogue Scale* ... 17
3.1.2. *Standard Gamble* ... 18
3.1.3. *Time Trade-Off* ... 19
3.1.4. Outros Métodos ... 20
3.1.4.1. *Person Trade-Off* (PTO) ... 20
3.1.4.2. *Ratio Scaling* (RS) ... 21
3.2. Medidas Genéricas Multi-Atributo Baseadas em Preferências ... 21
3.2.1. *Quality of Well-Being* ... 21
3.2.2. *Health Utilities Index* ... 22
3.2.3. *EuroQol-5D* ... 22
3.2.4. *Short-Form 6D* ... 23
3.3. Outros Métodos de Obtenção de Utilidades em Saúde ... 23
3.4. Estudos comparativos de métodos diretos de obtenção de utilidades ... 26
3.5. Medição de QVRS em Portugal ... 27
3.6. Influência da Numeracia e da Fluência na Expressão de Julgamentos de Valor ... 29
4. Metodologia Proposta ... 31
4.1. Desenho Metodológico ... 32
4.2. Abordagem MACBETH ... 35
4.2.1. Questionamento MACBETH e obtenção de Escalas de Valor ... 35
4.2.2. Programação Linear MACBETH adaptada ao Estudo ... 37
4.2.3. Abordagem MACBETH aplicada à Avaliação de Preferências em Saúde ... 38
4.3. Desenvolvimento de Questionário de Avaliação de Estados de Saúde ... 40

4.3.1. Selecção do Método Numérico ... 40
4.3.2. Seleção dos Estados de Saúde ... 41
4.3.3. Estruturação do Questionário ... 42
4.3.4. Implementação do Questionário *online* na Plataforma Qualtrics ... 43
4.3.4.1. Caraterização do Estado de Saúde Atual ... 43
4.3.4.2. Recolha de Julgamentos de Valor Numéricos – Time Trade-Off ... 43
4.3.4.3. Recolha de Julgamentos de Valor Não Numéricos – MACBETH ... 46
4.3.4.4. Avaliação das Preferências ... 47
4.3.4.5. Avaliação da Numeracia ... 48
4.3.5. Estratégia de Distribuição e População-Alvo ... 48
4.4. Obtenção de Escalas de Valor ... 48
4.4.1. Valoração Time Trade-Off ... 49
4.4.2. Valoração MACBETH ... 49
4.5. Critérios de Exclusão ... 50
4.6. Análise de Dados Recolhidos ... 51
4.6.1. Análise Estatística ... 52
4.6.2. Outras Análises ... 52
5. Resultados ... 54
5.1. Obtenção da Amostra Válida para análise ... 54
5.2. Caraterização da Amostra Populacional ... 58
5.2.1. Caraterização Sociodemográfica ... 58
5.2.2. Caraterização do Estado de Saúde Atual ... 60
5.2.3. Caraterização do Grau de Numeracia ... 61
5.3. Valoração dos Estados de Saúde ... 61
5.3.1. Comparação entre as Escalas ... 61
5.3.2. Correlação entre Escala MACBETH e Escala TTO ... 64
5.3.3. Ajuste da Escala MACBETH à Escala TTO ... 65
5.4. Caraterização das Preferências Manifestadas ... 66
5.5. Caraterização da Influência do Grau de Numeracia ... 67
5.5.1. Influência do Grau de Numeracia nas Preferências ... 68
5.5.2. Influência do Grau de Numeracia na Valoração de Estados de Saúde ... 69
5.5.3. Influência do Grau de Numeracia nas Exclusões ... 71
6. Discussão ... 72
6.1. Comparação entre MACBETH e TTO ... 72
6.2. Considerações relativamente à Numeracia ... 74
6.3. Comparação com Escala de Valores para o EQ-5D Portugal ... 75
6.4. Limitações da metodologia proposta ... 76
7. Conclusão ... 77
Referências ... 79
ANEXO A ... 84
ANEXO B ... 85

Lista de Figuras

Figura 1. Representação esquemática do conceito de QALY [38]. 11
Figura 2. *Standard Gamble* para um estado de saúde crónico i, considerado melhor do que a morte (adaptado de [8]). 18
Figura 3. *Time Trade-Off* para um estado de saúde crónico, considerado melhor do que a morte (adaptado de [8]). 19
Figura 4. Representação esquemática da metodologia proposta. As setas contínuas representam relações sequenciais e as setas a tracejado representam processos iterativos. 34
Figura 5. (a) Exemplo de matriz de julgamentos MACBETH consistente; (b) Escala proposta pelo *software* M-MACBETH (adaptado de [88]). 36
Figura 6. Representação esquemática dos diferentes passos necessários para avaliação de estados de saúde recorrendo ao método proposto, baseado na abordagem MACBETH. 39
Figura 7. Estruturação do questionário em cinco módulos, cada um correspondente à recolha de diferentes tipos de dados. 42
Figura 8. Exemplo de questão de avaliação se um dado estado de saúde é considerado melhor ou pior do que morte imediata. 45
Figura 9. Exemplo de questão de recolha de julgamentos para estados considerados melhores do que a morte. 45
Figura 10. Exemplo de questão de recolha de julgamentos para estados considerados piores do que a morte. 46
Figura 11. Exemplo de questão de recolha de julgamentos de diferença de atratividade com base na escala semântica MACBETH. 47
Figura 12. Aplicação do MACBETH para valoração do conjunto de estados de saúde apresentado ao Indivíduo 1. 50
Figura 13. Diagrama de fluxo da obtenção da amostra populacional válida para análise (G1, G2, G3, G4 – Conjunto de questionários obtidos relativos ao grupo de estados de saúde 1, 2, 3 e 4, respetivamente). 55
Figura 14. Descrição do número de exclusões obtidas tendo em conta a ordem de apresentação das tarefas de recolha de julgamentos de acordo com o método MACBETH e TTO. 57
Figura 15. Descrição do número de tarefas concluídas, nos questionários incompletos, tendo em conta a ordem de apresentação das tarefas de recolha de julgamentos de acordo com o método MACBETH e TTO. 57
Figura 16. Representação do valor médio obtido para cada estado, ordenados de acordo com a escala de valores MACBETH. A cor verde representa-se os valores obtidos recorrendo ao TTO e a cor azul os valores obtidos recorrendo ao MACBETH. 63
Figura 17. Valores médios atribuídos aos 25 estados de saúde pelo método TTO em função dos valores médios obtidos pelo método MACBETH. 64
Figura 18. Valores atribuídos a todos os estados de saúde pelo método TTO em função dos valores atribuídos pelo método MACBETH. 65

Figura 19. Representação dos valores obtidos recorrendo ao método TTO para a amostra total (ponto preto), para a amostra com elevada numeracia (ponto verde) e para a amostra com baixa numeracia (ponto azul). 70
Figura 20. Representação dos valores obtidos recorrendo ao método MACBETH para a amostra total (ponto preto), para a amostra com elevada numeracia (ponto verde) e para a amostra com baixa numeracia (ponto azul). 70
Figura 21. Sistema descritivo do EQ-5D (adaptado de [97]. 84

Lista de Tabelas

Tabela 1. Diferentes tipos de avaliação em saúde (adaptado de [31]). 7
Tabela 2. Estudos comparativos entre diferentes métodos diretos de obtenção de utilidades em saúde. 26
Tabela 3. Estudos de medição de QVRS baseada em preferência realizados em Portugal. 29
Tabela 4. Distribuição dos estados de saúde por grupo de acordo com o seu grau de severidade. ... 41
Tabela 5. Descrição, por critério, do número total de julgamentos excluídos associados a cada método (N = 243 questionários completos). 56
Tabela 6. Caraterização da amostra populacional do estudo antes e após exclusões. 58
Tabela 7. Distribuição das respostas dos indivíduos às dimensões do EQ-5D e ao EQ-5D VAS. 61
Tabela 8. Resultados obtidos relativamente à valoração do conjunto de 25 estados de saúde avaliados, recorrendo ao MACBETH e ao TTO. 62
Tabela 9. Caraterização do modo preferencial para expressão de julgamentos de valor manifestado pelos indivíduos da amostra. 66
Tabela 10. Principais comentários ao método baseado na abordagem MACBETH e ao método TTO. 66
Tabela 11. Caraterização das preferências manifestadas de acordo com a ordem de apresentação das tarefas de atribuição de julgamentos. 67
Tabela 12. Caraterização das preferências manifestadas relativamente ao modo de expressão de julgamentos de valor, de acordo com o grau de numeracia dos indivíduos. 68
Tabela 13. Correlação de Spearman entre escalas de valores obtidas recorrendo ao método MACBETH e TTO. 69
Tabela 14. Descrição do número de exclusões identificadas para a amostra com “Baixa numeracia” e “Elevada numeracia”. 71
Tabela 15. Coeficiente de correlação de Pearson entre as escalas de valores médios obtidas no presente estudo (TTO e MACBETH) e a escala definida para a população portuguesa. 76
Tabela 16. Caraterização da amostra populacional que avaliou cada um dos quatro conjuntos de estados de saúde definidos. 85

Lista de Abreviaturas

ATS: Avaliação de Tecnologias de Saúde

DALYs: *Disability Adjusted Life Years*

DCE: *Discrete Choice Experiments*

EQ-5D: *EuroQol-5D*

HTAi: *Health Technology Assessment international*

HUI: *Health Utilities Index*

HYE: *Healthy-years equivalent*

INAHTA: Rede Internacional de Agências de Avaliação de Tecnologias de Saúde

MACBETH*: Measuring Attractiveness by a Categorical Based Evaluation Technique*

NICE: *National Institute for Clinical Excellence*

PTO: *Person Trade-Off*

QALYs: *Quality Adjusted Life Years*

QVRS: Qualidade de Vida Relacionada com a Saúde

QWB: *Quality of Well-Being*

RICE: Rácio Incremental Custo-Efetividade

RS: *Ratio Scaling*

SF-6D: *Short-Form 6D*

SG: *Standard Gamble*

SiNATS: Sistema Nacional de Avaliação de Tecnologias de Saúde

SNS: Serviço Nacional de Saúde

TTO: *Time Trade-Off*

VAS: *Visual Analogue Scale*

1. Introdução

1.1. Motivação

Nos países desenvolvidos tem-se verificado, ao longo das últimas décadas, um aumento contínuo no desenvolvimento e introdução de novas tecnologias em saúde, nomeadamente de medicamentos, dispositivos médicos, procedimentos clínicos e cirúrgicos, entre outros. Esta situação surge como consequência do rápido desenvolvimento tecnológico, bem como da procura crescente de cuidados de saúde resultante do envelhecimento populacional, da maior prevalência de doenças crónicas e do aumento das expectativas dos pacientes em relação ao seu estado de saúde. [1] [2]

A rápida difusão de tecnologias em saúde permitiu à população beneficiar de melhorias na prestação de cuidados, traduzindo-se em ganhos significativos de saúde. No entanto, por outro lado, o aumento do número de intervenções médicas disponíveis, bem como, do custo a elas associado, constituiu um novo desafio para os decisores em saúde, uma vez que, os recursos financeiros, humanos e físicos disponíveis são limitados e não conseguem acompanhar este aumento. Na gestão de sistemas de saúde é, assim, necessário tomar decisões de alocação de recursos para uma determinada área/alternativa de ação em detrimento de outras, de modo a obter o máximo aproveitamento dos recursos disponíveis proporcionando à população cuidados de saúde que vão ao encontro das necessidades existentes e em linha com os objetivos institucionais, constrangimentos orçamentais e princípios como a equidade e o acesso universal. [3] [4]

O problema da alocação de recursos financeiros, particularmente, no que diz respeito à adoção de novas tecnologias em saúde tem sido alvo de um interesse crescente, tendo-se tornado particularmente urgente devido à atual situação de tensão derivada do rápido crescimento das despesas em saúde e do simultâneo abrandamento da economia. [2] Como consequência, tem-se vindo a desenvolver a Avaliação de Tecnologias de Saúde (ATS), uma área multidisciplinar, que pretende reunir dados para ajudar os decisores políticos, prestadores de cuidados de saúde e pacientes a entender o valor relativo das tecnologias. [5] [6]

A ATS recorre a diferentes metodologias de avaliação económica, que permitem avaliar as várias alternativas de ação, em termos dos ganhos de saúde obtidos e de custos de oportunidade. [7] Destaca-se, de entre as diferentes metodologias de avaliação económica, a análise custo-utilidade, um tipo de análise amplamente utilizada que torna possível a comparação entre intervenções de áreas diferentes da prestação de cuidados de saúde, através da medição dos ganhos de saúde obtidos numa medida generalizada, como por exemplo, os anos de vida ajustados pela qualidade (QALYs). [8]

Os QALYs são uma medida de benefícios de saúde que englobam não só o impacto das intervenções médicas no tempo de vida ganho pelo paciente mas também os resultados obtidos em termos de qualidade de vida relacionada com a saúde (QVRS). Nesta abordagem, cada ano de vida ganho é, assim, ponderado por um coeficiente que reflete a qualidade de vida que lhe está associada. [8] Estes coeficientes denominam-se utilidades e são calculados para os diferentes estados de saúde refletindo as preferências relativas de um indivíduo ou grupo de indivíduos.

Os métodos convencionais de elicitação de preferências em saúde incluem o *Visual Analogue Scale* (VAS), o *Time Trade-Off* (TTO) e o *Standard Gamble* (SG), não havendo, atualmente, um consenso em relação ao método a ser preferencialmente utilizado, uma vez que, cada método está associado a vantagens e desvantagens distintas. Tanto o SG como o TTO revelam uma elevada exigência cognitiva e as respostas obtidas podem ser influenciadas por fatores como aversão ao risco ou preferência temporal, por outro lado, o VAS é amplamente criticado por não apresentar uma fundamentação teórica evidente. [9]

A utilização de métodos quantitativos, isto é, métodos baseados na atribuição de julgamentos de valor numéricos para determinação das utilidades de saúde, tais como o VAS, o TTO e o SG, tem sido amplamente explorada e, para além, das desvantagens individuais associadas a cada método, estudos recentes determinam que competências, tais como a numeracia e a fluência, afetam a capacidade dos indivíduos em expressar julgamentos de valor em números ou palavras. [10] [11] [12]

O interesse crescente na medição da qualidade de vida relacionada com a saúde de forma sistematizada e recorrendo a um índice único baseado nas preferências dos indivíduos, associado a uma constatação das limitações e das dificuldades cognitivas dos métodos tradicionais de elicitação de preferências justifica o desenvolvimento de métodos alternativos para avaliação de estados de saúde e a realização de diferentes estudos de comparação entre valores obtidos recorrendo a diferentes métodos e de análise das dificuldades sentidas pelos indivíduos ao responder a diferentes protocolos. Surge, então, neste contexto a principal motivação para o desenvolvimento do presente trabalho.

1.2. Objetivos

O trabalho realizado nesta tese de mestrado apresenta como principal objetivo o desenvolvimento e teste de um novo método de avaliação de preferências em saúde baseado em julgamentos não numéricos recorrendo à abordagem MACBETH (*Measuring Attractiveness by a Categorical Based Evaluation Technique*). Numa primeira fase pretende-se desenhar um novo protocolo de avaliação de estados de saúde recorrendo à atribuição de julgamentos semânticos (não numéricos) de diferença de atratividade entre estados de saúde para obtenção de uma escala numérica de pontuações refletindo a atratividade relativa dos mesmos. Uma vez definido e estruturado o novo protocolo proposto, pretende-se implementá-lo simultaneamente com um método convencional de avaliação de preferências em saúde, amplamente utilizado pelos investigadores da área em diferentes estudos e baseado em julgamentos numéricos, com o intuito de tornar, assim, possível uma análise comparativa dos resultados obtidos recorrendo ao novo método proposto e ao método convencional. Para implementação de ambos os métodos referidos pretende-se explorar uma plataforma Web de construção e distribuição de questionários.

Adicionalmente pretende-se, também, avaliar as preferências dos indivíduos relativamente à expressão de julgamentos de valor em números ou através de categorias semânticas, ou seja, em relação aos dois métodos implementados.

Tendo em conta as questões anteriormente levantadas relativamente à influência de competências numéricas na valoração de estados de saúde, pretende-se, por último, caraterizar a numeracia dos indivíduos e analisar a influência desta competência na valoração dos estados de saúde e nas preferências manifestadas.

1.3. Estrutura do Documento

Para além deste capítulo introdutório esta tese encontra-se organizada em seis capítulos. No capítulo 2 é realizada uma breve descrição do âmbito da avaliação de tecnologias de saúde e são introduzidas as diferentes abordagens de avaliação económica. Neste capítulo é dado particular destaque às análises custo-utilidade, descrevendo-se a abordagem dos QALYs para determinação dos ganhos de saúde e os conceitos de utilidade e valor. Por fim, é feita uma breve referência à medição da qualidade de vida relacionada com a saúde. O capítulo 3 consiste numa revisão bibliográfica dos principais métodos de obtenção de utilidades em saúde encontrados na literatura, nomeadamente métodos diretos e instrumentos multi-atributo. São, também, discutidos neste capítulo alguns dos desafios encontrados na medição direta de utilidades em saúde e é dado destaque aos avanços na área realizados em Portugal. Por fim, é introduzida a questão da influência de competências como a numeracia e a fluência dos indivíduos na atribuição de julgamentos de valor. No capítulo 4 é apresentada a metodologia proposta de modo a atingir os objetivos definidos para este trabalho, descrevendo-se em detalhe as várias etapas que a constituem, entre as quais a elaboração de um questionário *online* para recolha de dados. No capítulo 5 são apresentados e analisados os principais resultados obtidos através da implementação da metodologia proposta. Por fim, o capítulo 6 e o capítulo 7 consistem, respetivamente, na discussão dos resultados obtidos e na apresentação das principais conclusões resultantes do trabalho desenvolvido.

2. Contexto

A importância da avaliação de diferentes intervenções em saúde, nomeadamente a avaliação de diferentes tecnologias, tem apresentado um interesse crescente ao longo das últimas décadas, uma vez que os seus resultados permitem informar e melhorar as decisões de alocação de recursos em todos os setores do sistema de saúde, quer públicos como privados. [13]

Este capítulo começa por apresentar uma breve descrição do âmbito da avaliação de tecnologias de saúde. De seguida, introduz-se a avaliação económica e os diferentes tipos de abordagens que podem ser seguidas, entre as quais se destacam as análises custo-efetividade, custo-utilidade e custo-benefício. De entre os diferentes tipos de avaliações económicas apresentadas é dada maior ênfase à análise custo-utilidade, abordando-se o uso dos QALYs para determinação dos ganhos de saúde e o conceito de utilidade subjacente a esta abordagem. Por último é feita uma breve referência à medição de qualidade de vida relacionada com a saúde e aos vários tipos instrumentos de medição existentes.

2.1. Avaliação de Tecnologias de Saúde

Nas últimas décadas, assistiu-se, nos países desenvolvidos, a uma vaga de inovação tecnológica com efeitos em diferentes áreas, nomeadamente na área da saúde, onde se refletiu sobretudo em alterações na prática clínica e na organização dos serviços de saúde. [14]

Embora o rápido desenvolvimento e introdução de novas tecnologias em saúde tenha permitido à população beneficiar de serviços de saúde melhorados, originando ganhos significativos de saúde, fez-se acompanhar por um aumento acentuado dos gastos o que constitui um novo desafio para os decisores em saúde, uma vez que os recursos disponíveis são limitados, nomeadamente os recursos financeiros.

Quando confrontados com decisões complexas de alocação de recursos os decisores em saúde têm tendência a tomar decisões baseadas em motivos para os quais não existe uma fundamentação sólida, tais como, padrões históricos e políticos, abordagens intuitivas ou regras *ad-hoc*. [3] [4] Este tipo de tomada de decisão não apresenta um conjunto de regras claras e tecnicamente bem sustentadas nem faz uma análise cuidadosa das várias alternativas de ação, identificando todos os custos e consequências a elas associados, conduzindo, deste modo, a um sob aproveitamento dos recursos disponíveis e levantando questões relativamente à sua transparência, estrutura e abrangência, uma vez que não incorpora explicitamente as preferências dos pacientes, as necessidades não atendidas e os valores éticos e sociais. [3]

Num contexto em que a saúde é um sector cada vez mais dependente da tecnologia e simultaneamente se verifica um abrandamento da economia conduzindo a uma necessidade acrescida de contenção de custos [2], a alocação de recursos tornou-se um problema central e é urgente o desenvolvimento de um conjunto de procedimentos sistematizados apropriados para a definição eficiente de prioridades de acordo com as necessidades crescentes da população e os recursos disponíveis. [15]

O campo da ATS surgiu, então, como resposta à crescente preocupação acerca da sustentabilidade económica do sistema e em relação à qualidade e variabilidade dos serviços prestados [16], sendo definido, de acordo com a Sociedade Internacional de Avaliação de Tecnologias de Saúde (HTAi), como um 'campo de pesquisa científica que pretende informar a tomada de decisão política e clínica acerca da introdução e difusão de tecnologias de saúde'. [17]

Na ATS são definidas como tecnologias de saúde todos os dispositivos e procedimentos médicos que possam ser utilizados por profissionais de saúde para prevenir, diagnosticar, tratar, acompanhar ou reabilitar o doente, tais como, produtos farmacêuticos, procedimentos médicos e cirúrgicos, equipamentos de tratamento e diagnóstico entre outras intervenções clínicas. [14] [17]

A ATS consiste, então, numa área multidisciplinar que pretende ajudar os decisores políticos, os prestadores de cuidados de saúde e também os pacientes e a população em geral a entender o valor relativo das tecnologias, recorrendo para tal a uma avaliação sistemática e transparente das propriedades e efeitos das tecnologias de saúde. [4] Esta avaliação assenta sobretudo na produção, síntese e/ou revisão sistemática de dados de evidência científica e não científica disponível sobre a segurança, eficácia, efetividade, custo e custo-efetividade de produtos ou serviços de saúde. [18] Além destes pontos são também, cada vez mais, considerados os fatores contextuais, tais como, impactos económicos, organizacionais, sociais e éticos que resultam da utilização das tecnologias em análise, bem como as alternativas disponíveis. [16] A ATS assenta, assim, não apenas sobre os aspetos que podem ser influenciados pelo uso de tecnologias de saúde mas também sobre os aspetos que as influenciam. [5]

Além de identificar as tecnologias mais efetivas e de maior valor e promover a sua utilização, a ATS permite também reduzir ou eliminar o uso de tecnologias que não sejam seguras ou efetivas, ou aquelas cujos custos são demasiado elevados comparados com os benefícios obtidos. [4] [5]

Desde a década de 1990 foram vários os países que têm vindo a alterar os seus métodos de tomada de decisão relativamente à alocação de recursos no sistema de saúde, tendo surgido várias agências governamentais e programas de ATS [4], bem como, cooperações e coordenações internacionais cujo principal objetivo é a identificação de tecnologias de maior valor relativo. [6] Entre estas destacam-se a Sociedade Internacional de Avaliação de Tecnologias de Saúde (HTAi) e a Rede Internacional de Agências de Avaliação de Tecnologias de Saúde (INAHTA). A HTAi é uma sociedade científica e profissional de produção e uso de ATS, tendo surgido em 2003 e sendo constituída, atualmente, por membros de mais de 65 países e de diferentes partes interessadas, incluindo investigadores, agências de ATS, decisores políticos, indústrias, universidades, prestadores de serviços de saúde e pacientes. [17] Esta sociedade define-se como um fórum neutro para colaboração, discussão e partilha de informação úteis e conhecimentos relacionados com ATS. [17] [19] Outra associação em destaque e que se relaciona com a HTAi é a INAHTA, que surgiu por volta de 1993 e é, atualmente, constituída por uma rede de 55 agências de ATS, ligadas aos governo regionais ou nacionais, de 32 países que cooperam entre si e compartilham informações sobre o seu trabalho de produção e divulgação de relatórios de ATS. [20] [21]

Uma das primeiras instituições a surgirem e a fazerem uso dos relatórios produzidos pelas agências de ATS, foi o *National Institute for Clinical Excellence* (NICE), em Inglaterra, cuja actividade

se centra em identificar, com base num processo de revisão sistemático e baseado na evidência, maioritariamente suportado pelo programa de ATS deste país, as tecnologias mais custo-efetivas e promover a sua implementação, fornecendo orientação a nível nacional, através do desenvolvimento de *guidelines*, na promoção de boa saúde e prevenção e tratamento de doenças. [6] [22] [23] O NICE serve, atualmente, como padrão a outros países, promovendo uma cultura de avaliação sistemática baseada em evidência de tecnologias de saúde. [4]

Embora, atualmente, a tendência comum a quase toda a Europa seja de um crescimento substancial de ATS e existam evidências de avaliações de elevada qualidade, esta é uma área ainda bastante desigual entre países e que enfrenta alguns desafios, nomeadamente ao nível da elaboração de análises cientificamente robustas e baseadas em evidências com relevância política. [13] [18]

Avaliação de Tecnologias de Saúde em Portugal

Atualmente, em Portugal, o INFARMED, I.P., é a entidade responsável pela avaliação de tecnologias de saúde, sendo que, desde há 15 anos que são realizadas avaliações para a tecnologia de saúde medicamento, nomeadamente avaliações económicas, utilizadas como instrumentos obrigatórios de apoio à decisão no âmbito de processos de comparticipação ou de avaliação prévia à sua aquisição pelos hospitais do Serviço Nacional de Saúde (SNS). [24] [25]

Seguindo a tendência de desenvolvimento da ATS, comum a quase toda a Europa, foi recentemente instituído, por esta entidade, o SiNATS – Sistema Nacional de Avaliação de Tecnologias de Saúde. [25] Este novo modelo que pretende incluir a integração de novas tecnologias de saúde, tais como, dispositivos médicos pretende, também, a introdução de medidas que assegurem a transparência, a previsibilidade, a equidade na utilização e a obtenção dos ganhos em saúde que justificaram o seu financiamento, avaliando-se, deste modo, o valor das tecnologias ao longo do seu ciclo de vida. Pretende-se, adicionalmente, um aprofundamento do envolvimento de Portugal nos exercícios comunitários de criação de um sistema integrado de avaliação de tecnologias de saúde. [25]

Com a implementação do SiNATS, o INFARMED, I.P., quer contribuir para a maximização dos ganhos em saúde e da qualidade de vida dos cidadãos garantindo que o SNS investe da melhor forma possível os seus recursos, assegurando a sustentabilidade do sistema e a utilização eficiente dos recursos públicos. [25]

2.2. Avaliação Económica em Saúde

Nos últimos anos, com base na premissa de limitação de recursos financeiros na área da saúde e perante uma consciencialização crescente da necessidade de utilização eficiente dos recursos disponíveis, de modo a maximizar os benefícios obtidos pela população a nível de saúde, tornou-se necessário fazer escolhas não apenas baseadas nas necessidades da população, e em serviços e intervenções clinicamente efetivos mas também que funcionem a menor custo, isto é, que sejam mais

custo-efetivos. [26] As metodologias baseadas em princípios económicos foram, assim, introduzidas na área da saúde pelos economistas trazendo a variável custo diretamente para a decisão e tornando o constrangimento de recursos financeiros explícito. [26]

A Economia é, simplificando, a ciência da escassez de recursos [27] e do estudo de diferentes alternativas de alocação dos mesmos [28], apresentando uma avaliação económica duas características essenciais: lida com a comparação entre *inputs* e *outputs* das alternativas em estudo, também designados como custos e consequências, e preocupa-se com as escolhas procurando identificar e tornar explícito um conjunto de critérios que podem ser úteis na tomada de decisão entre diferentes usos dos recursos disponíveis. [15]

Em termos gerais uma avaliação económica define-se como uma "análise comparativa de cursos alternativos de ação em termos de seus custos e consequências". [15] Assim, a tarefa básica de qualquer avaliação económica é identificar, medir, atribuir valor e comparar os custos e consequências das alternativas de ação a serem consideradas, podendo uma das alternativas corresponder a não se realizar nenhum tipo de ação. [15] [29]

Existem diferentes metodologias de avaliação económica de tecnologias e programas de saúde comumente encontradas em literatura, sendo apenas consideradas avaliações completas, de acordo com a Tabela 1, aquelas que consideram a comparação de duas ou mais alternativas em termos dos seus custos e benefícios. Entre estas destacam-se as análises custo-benefício, as análises custo-efetividade e as análises custo-utilidade que, embora partilhem entre si o objetivo chave de minimização dos custos e maximização dos benefícios obtidos, isto é, dos ganhos em saúde, diferem sobretudo na forma de medição e valorização destes, considerando diferentes tipos de informação, de acordo com as descrições apresentadas nas secções seguintes. [15] A escolha acerca de qual a melhor técnica a adotar em avaliações económicas deve refletir as diferentes situações de tomada de decisão e as informações que se pretendem analisar. [30]

Tabela 1. Diferentes tipos de avaliação em saúde (adaptado de [31]).

<table>
<tr><th colspan="2" rowspan="3"></th><th colspan="3">Análise de custos e consequências?</th></tr>
<tr><th colspan="2">Não</th><th rowspan="2">Sim</th></tr>
<tr><th>Só consequências</th><th>Só custos</th></tr>
<tr><td rowspan="4">Comparação de duas ou mais alternativas?</td><td rowspan="2">Não</td><td colspan="2">Avaliação parcial</td><td>Avaliação parcial</td></tr>
<tr><td>• Descrição de resultados</td><td>• Descrição de custos</td><td>• Descrição de custos e resultados de um só programa</td></tr>
<tr><td rowspan="2">Sim</td><td colspan="2">Avaliação parcial</td><td>Avaliação económica completa</td></tr>
<tr><td>• Avaliação da Eficácia ou Efetividade</td><td>• Análise de Custos</td><td>• Minimização de Custos
• Análise Custo-Efetividade
• Análise Custo-Benefício
• Análise Custo-Utilidade</td></tr>
</table>

Relativamente aos custos considerados em avaliações económicas em saúde estes abrangem os custos diretos, isto é, custos fixos e variáveis diretamente associados à prestação do serviço de

saúde, e os custos indiretos que, consistem, por exemplo em custos associados à diminuição de produtividade do doente e dos respetivos familiares. É ainda possível identificar custos intangíveis que incluem aspetos como o valor monetário da dor, ansiedade ou sofrimento do doente e da família e que são de difícil mensuração. [4] [31]

Embora a avaliação económica apresente o grande potencial de comparar diretamente diferentes alternativas de ação [32] e seja considerada, atualmente, uma das componentes vitais da ATS, apresenta também algumas limitações, nomeadamente relativamente ao tempo e aos custos envolvidos em cada análise, à enumeração e estimativa de custos e benefícios sendo, por vezes, a obtenção destes dados um processo difícil e variável consoante a perspetiva que se assume (da sociedade, dos pacientes, dos prestadores de cuidados ou dos decisores políticos [33]), e ainda o fato de não considerar explicitamente o orçamento disponível. [7] [32] Tendo em conta, algumas destas limitações é por vezes difícil estabelecer a ligação entre as análises económicas realizadas e a tomada de decisão. [7]

2.2.1. Análise Custo-Benefício

Na análise custo-benefício as consequências das diferentes alternativas em análise, ou seja, os benefícios de saúde obtidos, bem como, os custos associados a cada uma das alternativas são expressos em unidades monetárias, como por exemplo, o euro. [15] Uma vez que todos os benefícios são convertidos em unidades monetárias, podem ser considerados efeitos únicos ou múltiplos e que não sejam necessariamente comuns a todas as alternativas em estudo. [14]

Neste tipo de análise os resultados podem ser apresentados sob a forma de um rácio entre custos e benefícios, ou como uma soma que representa o benefício líquido de uma alternativa sobre as outras (cujo resultado pode também ser negativo). [4] Desta forma, é possível para o decisor comparar diretamente as alternativas em estudo de modo a encontrar qual ou quais maximizam a razão entre custos e benefícios ou o benefício líquido obtido. [15] [34]

Esta técnica é, então, considerada o método de avaliação económica mais abrangente, uma vez que os benefícios são medidos em unidades monetárias e, como tal, os resultados obtidos não são restritos à área da saúde, permitindo ao decisor não só tomar conhecimento do retorno obtido através de investimentos em saúde mas também compará-los com investimentos noutras áreas da economia. [35]

No entanto, este tipo de análise apresenta como grande desvantagem a dificuldade que existe em medir e quantificar em unidades monetárias diferentes benefícios de saúde, por exemplo, maior qualidade de vida para o doente ou aumento da sua satisfação, bem como a própria vida. [15] A atribuição de valor monetário a diferentes estados de saúde e a benefícios muitas vezes intangíveis é um assunto bastante discutido sendo a principal razão para que esta técnica não seja amplamente aceite em economia de saúde. [36] É, também, discutido o fato de este tipo de análise poder discriminar aqueles com menor probabilidade de virem a ser produtivos no futuro, como por exemplo, os idosos ou os desempregados. [27]

Apesar da existência de alguns pontos de discussão, esta metodologia de avaliação económica destaca-se por permitir a comparação entre investimentos na área da saúde com investimentos noutras áreas, sendo por isso, bastante utilizada quando esse é o objetivo pretendido. [14]

2.2.2. Análise Custo-Efetividade

A análise custo-efetividade é o tipo de análise de avaliação económica mais encontrado na literatura de economia de saúde, sendo frequentemente aplicado na avaliação de medicamentos, avaliação de tecnologias ou planeamento de programas de saúde. [14] [27]

Este método consiste na análise comparativa de custos e consequências de duas ou mais alternativas de ação, onde os custos são quantificados em unidades monetárias e os benefícios de saúde são quantificados através de unidades naturais adequadas ao problema em análise. [15] Como unidades naturais, na área da saúde, incluem-se diferentes resultados clínicos como, por exemplo, os anos de vida ganhos, o número de vidas salvas, as complicações evitadas ou os casos diagnosticados. [35]

Este tipo de análise é muito útil em situações em que o decisor apresenta um determinado orçamento financeiro, cuja alocação pretende otimizar, e um conjunto limitado de opções numa dada área de saúde, sendo que, os benefícios de todas as opções apresentam um efeito de saúde comum, podendo este ser atingido em diferentes graus pelas diferentes opções. [15] É o caso, por exemplo, de dois tipos de medicação alternativas, ambas com efeito ao nível do aumento da esperança de vida, sendo que, uma medicação pode ter como consequência o aumento da esperança de vida num maior número de anos do que a outra. [36] O principal objetivo desta análise é, então, fornecer informações sobre a eficiência relativa de alternativas de ação que servem o mesmo objetivo, identificando as que permitem alcançar maior benefício para a população por unidade de investimento. [14] [29]

Na análise custo-efetividade a apresentação dos resultados quando as alternativas em análise são mutuamente exclusivas, por exemplo, quando se pretende comparar uma nova tecnologia com a tecnologia padrão habitualmente utilizada, sendo que, apenas uma delas pode ser escolhida, é feita recorrendo a uma comparação entre as diferenças de custos e as diferenças de benefícios obtidos entre as alternativas, na forma de razão entre a diferença de custos e a diferença de benefícios obtendo-se um quociente a que se chama rácio incremental custo-efetividade (RICE). [14] Este rácio é o resultado mais comum das análises custo-efetividade permitindo conhecer o custo adicional de uma alternativa em comparação com os benefícios adicionais da mesma, sendo que, quanto menor o valor do RICE mais custo-efetiva é alternativa considerada. [14] [37]

Em comparação com a análise custo-benefício, a análise custo-efetividade, apresenta como vantagem a utilização de unidades naturais para quantificar a efetividade não sendo necessário atribuir aos benefícios de saúde valor monetário, no entanto, existem por vezes dificuldades associadas à obtenção de bons dados de efetividade, bem como à definição dos resultados clínicos de efetividade mais importantes para reportar pois esta técnica não permite considerar múltiplos efeitos. [15] [35] Esta metodologia apresenta, ainda, outras desvantagens como o fato de não permitir

a comparação entre alternativas que não apresentem efeitos clínicos comuns, não sendo possível, também, a comparação de intervenções em saúde com intervenções em outras áreas. [36]

2.2.3. Análise Custo-Utilidade

A análise custo-utilidade é um método de avaliação económica que apresenta em alguns aspetos características semelhantes à análise custo-efetividade, sendo até considerada por alguns autores um caso particular desta. [15] [35] Estas metodologias diferem entre si, sobretudo, na forma de avaliação das consequências das diferentes alternativas em análise, sendo que, no caso da análise custo-utilidade os benefícios das várias alternativas são quantificados recorrendo a valores de utilidade que têm em conta as preferências individuais dos pacientes, ou da população em geral, por um determinado estado de saúde caracterizado por condições específicas. [14]

Geralmente, neste tipo de análise a medida utilizada para avaliar os benefícios obtidos são os QALYs (Secção 2.3.), que permitirem incorporar o valor da qualidade de vida (morbilidade) ganha pelos pacientes e também os anos de vida ganhos (mortalidade). Assim, enquanto numa determinada análise custo-efetividade os benefícios poderiam ser valorizados em termos de anos de vida ganhos, na análise custo-utilidade é valorizada também a qualidade dos anos de vida ganhos.

Tal como nas análises custo-efetividade, os resultados das análises custo-utilidade são expressos através de um RICE (razão entre a diferença de custos entre duas alternativas e a diferença de benefícios obtidos pelas mesmas) que representa, geralmente, o custo de obtenção de um QALY adicional. Quando se procede à comparação entre diferentes alternativas de ação, as que apresentem os valores de custo por QALY ganhos mais baixos deverão ser escolhidas. [38] Existem, no entanto, outras medidas utilizadas, tais como, o custo por anos de vida ajustados pela incapacidade (DALYs) ou o custo por anos equivalentes saudáveis (HYE). [15]

Estes resultados permitem incorporar consequências singulares e múltiplas e que não são necessariamente comuns a todas as alternativas em análise, ao contrário da análise custo-efetividade em que cada análise avalia uma consequência comum às várias alternativas em estudo. Assim, outra vantagem associada à análise custo-utilidade advém da possibilidade de incluir, na medida dos benefícios, uma extensa gama de resultados relevantes, bem como, comparar alternativas aplicadas em diferentes áreas de saúde, ou seja, sem um resultado clínico comum entre todas. [14]

No entanto, esta metodologia também apresenta algumas limitações, nomeadamente, o fato de apenas avaliar benefícios do ponto de vista de ganhos em saúde não tendo em conta benefícios sociais e apresenta ainda alguns problemas metodológicos associados ao uso de medidas de valorização dos benefícios obtidos dependentes de utilidades para as quais não existe um método de cálculo *standard* usado universalmente. [32]

Apesar de algumas limitações, as análises custo-utilidade têm-se tornado cada vez mais populares constituindo uma área de estudo em crescimento que tem abordado um conjunto diversificado de temas e tem apresentado estudos cada vez mais rigorosos e de acordo com as práticas metodológicas recomendadas. [39] Este é o tipo de abordagem de avaliação económica recomendada pelo NICE na avaliação de tecnologias de saúde. [23] [40]

2.3. Medição de Ganhos em Saúde através do uso de QALYs

2.3.1. Conceito de QALY

Os QALYs são, como referido anteriormente, uma medida de ganhos em saúde amplamente utilizada no auxílio à tomada de decisões relativamente à alocação de recursos quando se recorrem a análises custo-utilidade. [41] Este conceito surgiu, pela primeira vez, na década de 70 sendo, originalmente, desenvolvido como uma medida de efetividade para as análises custo-efetividade que pretendia incorporar simultaneamente numa única medida de ganhos em saúde efeitos de morbilidade (ganhos em qualidade de vida) e efeitos de mortalidade (ganhos em quantidade de vida). [15] [41]

O conceito de QALY, medida única de ganhos de saúde que combina os efeitos de uma intervenção em termos de anos de vida ganhos e da qualidade de vida experienciada durante esses anos, é apresentado na Figura 1, onde é apresentada uma situação hipotética de comparação entre duas alternativas de ação (com e sem programa ou intervenção de saúde). [38] De acordo com a Figura 1, um doente não sujeito a qualquer programa vive menos anos, acabando por morrer no momento designado por Morte e apresenta uma qualidade de vida relacionada com a saúde que se deteriora de acordo com a curva 'Sem programa'. Com o programa o indivíduo vive mais tempo, morrendo no momento designado por Morte', apresentando uma qualidade de vida que se deteriora mais lentamente, de acordo com a curva 'Com programa'. [35] A área entre as duas curvas, isto é, a diferença entre os QALYs gerados com e sem programa, corresponde ao número de QALYs ganhos com programa em estudo. [35]

A área entre as duas curvas pode ser dividida em duas zonas: a zona A e a zona B. A zona A corresponde ao número de QALYs ganhos em qualidade, isto é, ganhos devido a melhorias na qualidade de vida relacionada com a saúde do indivíduo durante o tempo de vida que este teria se não fosse submetido ao programa em estudo. A zona B corresponde ao número de QALYs ganhos devido ao prolongamento da vida associado ao programa. [38]

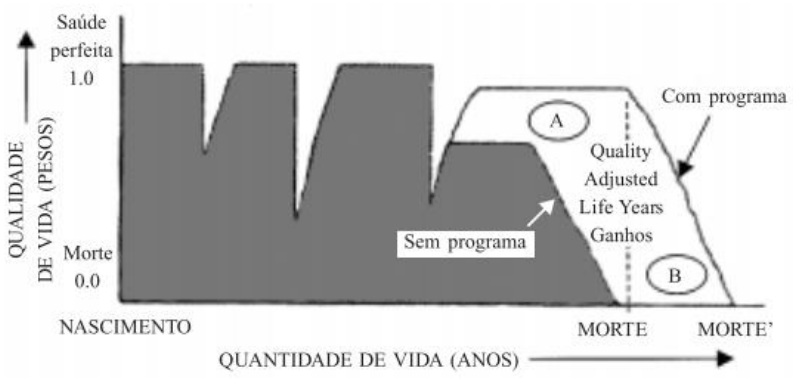

Figura 1. Representação esquemática do conceito de QALY [38].

No entanto, nem todos os programas ou intervenções de saúde apresentam este tipo de curvas, podendo haver casos em que a análise é mais complicada. É o caso, por exemplo, de tratamentos de

cancro que provocam uma diminuição da qualidade de vida a curto prazo mas permitem obter ganhos em QALYs a longo prazo ou determinadas intervenções/tratamentos que embora apresentem efeitos positivos a curto prazo provocam efeitos secundários a longo prazo. [38]

2.3.2. Cálculo dos QALYs

Como referido anteriormente, na abordagem dos QALYs cada ano de vida ganho é ponderado pela qualidade de vida relacionada com a saúde experienciada nesse ano [42] Os ajustes da qualidade de vida são baseados num conjunto de valores ou pesos de preferências chamados de utilidades, que reflectem as preferências relativas dos pacientes, ou da população em geral, pelos diferentes estados de saúde possíveis (Secção 2.4.). [15] Em economia de saúde, as utilidades são medidas numa escala cardinal intervalar entre 0 e 1, onde o valor 0 representa a morte e o valor 1 representa um estado de saúde perfeita sendo aos restantes estados de saúde atribuídos valores entre 0 e 1 de acordo com métodos adequados de medição de preferências em saúde. [8] [15] Realça-se que esta escala não exclui estados de saúde piores que a morte, sendo a estes atribuídos valores negativos. [43] Assim, um ano de vida de saúde perfeita corresponde a um QALY, enquanto, um ano de vida com uma utilidade inferior a 1, por exemplo x, corresponde a x QALYs. [38] [42]

Para determinar o valor corrente de QALYs que um indivíduo espera ganhar no futuro faz-se a atualização do valor de um ano de vida futura para um valor corrente por aplicação de uma taxa de desconto r, de modo a incorporar a ideia de que os indivíduos preferem ganhos em saúde no presente ao invés de ganhos no futuro. [38] A recomendação feita pelo NICE é que a taxa de desconto seja de 3.5% por ano, contudo, este é ainda um assunto em discussão na área. [8]

2.3.3. Limitações

A medição de ganhos em saúde recorrendo ao uso de QALYs apresenta uma grande atratividade por ser, em teoria, uma medida única, facilmente compreensível, que capta diferentes ganhos em saúde, em termos de qualidade e quantidade de vida, independentemente da doença e do tipo de paciente, permitindo assim o estabelecimento de uma medida padrão que possibilita a comparação entre diferentes tipos de programas, intervenções ou tecnologias de saúde, e consequente direcionamento dos recursos disponíveis para as alternativas que apresentam potencial para gerar mais QALYs. [15] [41] [38]

No entanto, esta abordagem apresenta também algumas limitações, nomeadamente todos os QALYs têm igual valor social, isto é, esta abordagem não tem em conta condições de saúde e características pessoais dos indivíduos, tais como, idade, sexo, gravidade da doença ou papel social do indivíduo. [8] O uso de QALYs pressupõe ainda a aditividade temporal das utilidades, isto é, assume-se que os valores de utilidade dos diferentes estados de saúde não são afetados pelos estados anteriores ou posteriores nem pelo tempo decorrido em cada estado. [30] Além destas limitações, um dos pontos de discussão relativamente à utilização desta metodologia recaí sobre a utilização de utilidades para ponderar a qualidade de vida relacionada com a saúde associada a

diferentes estados de saúde, sendo que a obtenção destas utilidades requer recursos extra e o método ao qual se recorre não é especificado no conceito de QALY, podendo as utilidades diferir ao nível das preferências que captam (governadores, gestores/administradores, população em geral ou qualquer outro grupo ou indivíduo), bem como nos princípios subjacentes ao método utilizado para a sua determinação, uma vez que, existem diferentes metodologias de avaliação de preferências com vantagens e desvantagens distintas. [38]

Em conclusão, embora a abordagem dos QALYs permita avaliar os benefícios em saúde obtidos a partir de diferentes alternativas de ação e, como tal, seja bastante útil na informação da tomada de decisão em saúde no que diz respeito à maximização dos benefícios de saúde obtidos, esta abordagem não inclui outros objetivos, tais como, equidade, justiça e objetivos políticos. [41]

2.4. Conceito de Utilidade, Valor e Preferência

A medida comumente utilizada para refletir as intensidades das preferências dos indivíduos, por diferentes estados de saúde é denominada de utilidade [15], podendo estas ser de dois tipos: ordinais e cardinais. [38]

As utilidades ordinais dizem respeito a uma escala que ordena por ordem de preferência os diferentes resultados de saúde, sendo estas facilmente obtidas através do estabelecimento por parte dos indivíduos de uma ordenação dos estados de saúde de acordo com as suas preferências. [38] Por outro lado, as utilidades cardinais correspondem a um conjunto de valores numéricos atribuídos aos diferentes estados de saúde, representando o seu peso de preferência. [15] As utilidades cardinais são, geralmente, definidas numa escala cardinal intervalar, habitualmente estabelecida usando como âncoras 0 e 1, onde o estado de saúde perfeita (ou melhor estado de saúde imaginável) é associado ao valor de utilidade 1 e a morte ao valor de utilidade 0, apresentando os restantes estados maior valor de utilidade quanto mais desejáveis forem. Estados considerados pior do que a morte também são possíveis assumindo valores de utilidade negativos. [43] O fato de a escala utilizada ser intervalar significa que um intervalo de igual comprimento tem a mesma interpretação, independentemente da parte da escala, por exemplo, uma alteração na saúde de 0.2 para 0.3 é equivalente a uma alteração de 0.8 para 0.9. [8]

Embora alguns autores não distingam a utilização do termo preferência, utilidade e valor, estes representam conceitos diferentes. Assim, o termo preferência é o termo mais abrangente que inclui o conceito subjacente tanto a utilidade como a valor, sendo utilidade e valor dois tipos diferentes de preferência. [15] O conceito de utilidade e valor diferem no que diz respeito às técnicas utilizadas na medição das preferências, nomeadamente na forma como se descrevem e formulam as questões durante o processo de obtenção destas. [4] [15] As utilidades representam, assim, a intensidade das preferências dos indivíduos em relação a determinados resultados, quando o processo de medição envolve uma situação de incerteza e risco, descritos com recurso a probabilidades. Por outro lado, os valores representam a intensidade das preferências quando o método de medição não envolve riscos ou incertezas. [15] [44]

2.5. Medição da Qualidade de Vida Relacionada com a Saúde

Nos últimos anos, decisores políticos, prestadores de cuidados de saúde e pacientes têm manifestado interesse num maior envolvimento dos pacientes e do público em geral na tomada de decisões em saúde, pretendendo-se, deste modo, que as decisões tomadas vão, cada vez mais, ao encontro das necessidades e aspirações dos pacientes e estejam de acordo com as perspetivas da população em geral. [45]

Neste contexto passou-se, assim, a dar maior importância à medição da qualidade de vida relacionada com a saúde, uma medida de resultados reportada pelo paciente, que incorpora a sua perceção de doença correlacionando-se com a sua capacidade funcional e bem-estar. [9] [46]

A qualidade de vida relacionada com a saúde é um conceito subjetivo e multidimensional de difícil definição que inclui todos os aspetos da qualidade de vida de alguma forma relacionados com a saúde, tais como, o bem-estar físico, funcional, social/ocupacional e emocional/psicológico do indivíduo. [47] [48] A sua medição pretende, assim, inferir acerca dos efeitos positivos e negativos de uma dada condição e do seu tratamento do ponto de vista do paciente, combinando esses efeitos numa medida composta. [48]

Instrumentos de Medição da QVRS

Têm sido vários os instrumentos desenvolvidos para medição da qualidade de vida relacionada com a saúde, habitualmente, na forma de questionários de auto-preenchimento (unidimensionais ou multidimensionais) ou protocolos de entrevistas semiestruturadas. [47] Estes instrumentos podem ser classificados em três grupos: instrumentos específicos, perfis de saúde genéricos e medidas baseadas em preferências. [46] [48]

Os instrumentos específicos, como o próprio nome indica, foram desenvolvidos para medição da qualidade de vida relacionada com a saúde focada em doenças específicas, populações específicas ou para avaliação de funções ou problemas específicos. [48] Como exemplo refere-se o instrumento *Catquest*, desenvolvido com o objetivo específico de analisar resultados da cirurgia às cataratas. [49]

Estes instrumentos são, habitualmente, utilizados quando se pretende estabelecer comparações entre indivíduos com características idênticas, sendo estas características as únicas relevantes na definição do resultado pretendido. Este tipo de instrumento apresenta a vantagem de se relacionar intimamente com as áreas rotineiramente analisadas pelos médicos e, uma vez que são específicos para um determinado tipo de paciente em estudo têm uma maior probabilidade de resposta visto que apenas incluem aspetos considerados relevantes para estes. [46]

Por outro lado, os instrumentos genéricos, tais como, os perfis de saúde genéricos e as medidas baseadas em preferências, permitem a obtenção de valores para o estado de saúde genérico dos indivíduos e são aplicáveis a vários pacientes com vários tipos de doenças, apresentando como maior vantagem o fato de poderem ser aplicados em diferentes áreas e em diferentes populações independentemente da condição em estudo, permitindo uma comparação do impacto relativo de

diferentes programas de saúde na QVRS. Estes instrumentos podem, no entanto, não captar alterações em condições específicas. [46]

Os perfis genéricos de saúde consistem num conjunto de pontuações relativas às várias dimensões da QVRS consideradas e, normalmente, estas não são agregadas numa pontuação global única. Como exemplo deste tipo de instrumento genérico destaca-se o SF-36 [50].

Os instrumentos de medição da QVRS baseados em preferências, por outro lado, dão origem a um índice único de pontuação global que resulta da agregação das diferentes dimensões através de um sistema em que se obtêm o peso relativo de cada dimensão. [38] A pontuação final obtida é baseada na medição das preferências dos indivíduos relativamente aos diferentes resultados produzidos ou advertidos por uma intervenção ou programa de saúde. [48] [51] Diferentes instrumentos de medição da QVRS baseados em preferências são discutidos na Secção 3.2.

Embora ambos os instrumentos genéricos ou específicos possam, assim, dar origem quer a perfis de saúde quer a índices, apenas os índices obtidos através de instrumentos de medição baseados em preferências, numa escala intervalar incluindo a saúde perfeita e a morte, apresentam potencial de utilização enquanto fatores de ajuste da QVRS, ou seja, traduzem as utilidades dos estados de saúde e permitem o cálculo dos QALYs, nas análises custo-utilidade. [48]

Neste capítulo foi contextualizada e descrita a abordagem dos QALYs para medição de ganhos em saúde, nomeadamente nas análises custo-utilidade que desenrolam, cada vez mais, um papel fundamental na ATS. Foram, ainda, introduzidos os conceitos de utilidade, valor e medição da QVRS subjacentes a esta abordagem.

Tendo em conta o principal objetivo deste trabalho que, consiste no desenvolvimento de um novo protocolo de valoração de estados de saúde, baseado em julgamentos não numéricos, de acordo com a abordagem MACBETH e na sua implementação conjuntamente com um método convencional, foi efetuada uma revisão bibliográfica com o intuito de identificar os principais métodos, encontrados na literatura, de valoração de estados de saúde e analisar as suas principais limitações teóricas e empíricas, bem como, a influência de competências numéricas e verbais dos indivíduos na expressão de julgamentos de valor. A revisão bibliográfica efetuada é apresentada no capítulo seguinte.

3. Revisão Bibliográfica

Para que se possam estimar os QALYs resultantes de um programa ou intervenção de saúde, é necessário a medição da qualidade de vida relacionada com a saúde associada aos resultados obtidos, recorrendo, para tal, a instrumentos de medição baseados em preferências de modo a obter índices que refletem utilidades cardinais.

Neste capítulo é, então, apresentada uma revisão bibliográfica com o intuito de identificar e discutir os diferentes métodos de medição de utilidades em saúde documentados na literatura, sendo que estes se distinguem em dois tipos: os métodos diretos e os métodos indiretos. [48]

Assim, na primeira secção são abordados os métodos utilizados para medir diretamente as preferências dos indivíduos por diferentes resultados de saúde, entre os quais o *Visual Analogue Scale* (VAS), o *Time Trade-Off* (TTO) e o *Standard Gamble* (SG). Enquanto na secção seguinte é feita referência a alguns métodos indiretos, também denominados de sistemas multi-atributo de classificação de estados de saúde baseados em preferências, uma vez que fazem uso de sistemas multi-atributo de classificação de estados de saúde, permitindo obter utilidades através do recurso a fórmulas de pontuação adequadas. Entre estes, destacam-se o *Quality of Well-Being* (QWB), o *Health Utilities Index* (HUI), o *EuroQol-5D* (EQ-5D) e o *Short-Form 6D* (SF-6D).

No presente capítulo é, também, dado destaque a novos métodos que têm vindo a ser desenvolvidos de modo a colmatar algumas das limitações identificadas nos métodos convencionais, fazendo-se, também referência a estudos, na área da saúde, de comparação entre métodos diretos de obtenção de utilidades e discutindo-se alguns dos desafios encontrados.

Por último são enunciados os avanços na área realizados em Portugal e é apresentada uma revisão acerca da influência de competências como a numeracia e a fluência dos indivíduos na obtenção de utilidades.

O ponto de partida para este capítulo de revisão bibliográfica foi uma pesquisa exaustiva pelas principais bases de dados com relevância na área - PubMed, ScienceDirect, Wiley Online Library e Google Scholar - recorrendo-se a palavras-chave como *CUA economic evalutation*, QALYs, *quality of life, utilities, measuring preferences, generic preference measures, MCDM, DCE, numeracy, value elicitation* combinadas entre si e com *Health* e Portugal.

3.1. Instrumentos Diretos de Obtenção de Utilidades

A medição de utilidades em saúde é um assunto complexo, existindo diferentes formas de determinar preferências relativamente a diferentes estados de saúde. Os analistas podem, então, adotar três estratégias diferentes de obtenção de valores de utilidades, que consistem na utilização de julgamentos de especialistas, obtenção de valores a partir de literatura relevante, ou recorrer a métodos de medição direta. [38] [52]

As duas primeiras fontes de obtenção de utilidades são potenciais fontes de erros, por exemplo, os julgamentos feitos pelos especialistas podem ser errados ou a literatura consultada pode ser inapropriada, não coincidindo os estados de saúde avaliados nesta com os que se pretendem avaliar,

sendo os indivíduos do estudo inapropriados ao novo estudo a ser realizado ou os instrumentos de medida utilizados não serem credíveis. [38]

Assim, a medição direta de utilidades é a estratégia comumente seguida, não existindo, no entanto, um consenso relativamente ao método mais adequada a esta medição, admitindo-se a importância da escolha deste de acordo com o problema em análise, uma vez que, diferentes métodos dão origem à obtenção de diferentes valores. [52] [53]

Os métodos diretos utilizados regularmente para elicitação de preferências são o *Visual Analogue Scale*, o *Standard Gamble* e o *Time Trade-Off*, que passamos a descrever.

3.1.1. *Visual Analogue Scale*

O método *Visual Analogue Scale* consiste numa escala de pontuação baseada na teoria psicométrica, em que é pedido aos indivíduos que avaliem um conjunto de estados de saúde de acordo com as suas preferências, sem ser necessária a atribuição direta de valores numéricos. [54] [55]

Este é o método mais simples de avaliação de preferências cardinais em saúde [8], consistindo apenas na definição de uma escala representada por uma linha reta, com habitualmente 10 cm, onde no topo da escala é colocado o estado de saúde com maior preferência, normalmente associado ao estado “melhor saúde imaginável” e, no final da escala é colocado o estado com menor preferência, geralmente “pior saúde imaginável” ou “morte”. [15] Depois de estabelecida a escala com os dois extremos bem definidos, é pedido aos indivíduos que introduzam os restantes estados de saúde, de acordo, com a sua ordem de preferência e de modo a que os intervalos entre os diferentes estados de saúde correspondam às diferenças percebidas em termos de preferência. [8]

As linhas que representam a escala podem variar em comprimento, ser verticais ou horizontais e apresentar ou não intervalos marcados com diferentes números. [54]

Apesar de a sua simplicidade ser bastante apelativa este método é, geralmente, apenas utilizado como um exercício de introdução a outros métodos, permitindo aos indivíduos familiarizar-se com o conceito de comparação entre estados de saúde e obter a sua correcta ordenação por preferência. [48] [56] A sua utilização como método de medição de preferências é limitada devido, sobretudo, à sua fraca fundamentação económica, uma vez que, não envolve uma tarefa de escolha ou decisões individuais sob incerteza mas, apenas, uma tarefa de pontuação. [56] [55] Este método apresenta, também, viés de escala, nomeadamente de fim de escala, devido aos indivíduos serem avessos a colocar estados de saúde próximos dos extremos. [8]

Outras abordagens semelhantes ao VAS são por vezes utilizadas, variando apenas o formato de apresentação da escala utilizada. Como exemplo, destaca-se a escala de pontuações categórica que apresenta diferentes categorias, por exemplo entre 0 e 10, igualmente espaçadas.

3.1.2. *Standard Gamble*

O *Standard Gamble* é o método clássico de medição de utilidades em saúde, baseado na teoria da Utilidade Esperada desenvolvida por von Neuman e Morgensten, em que os indivíduos são confrontados com uma escolha entre duas alternativas, sendo que, uma delas envolve um elemento de incerteza e risco. [8] [54] Para estados de saúde crónicos considerados melhores do que a morte, a tarefa de escolha apresentada é entre permanecer com certeza num dado estado de saúde ou arriscar uma aposta onde os resultados possíveis são saúde perfeita ou morte. [8]

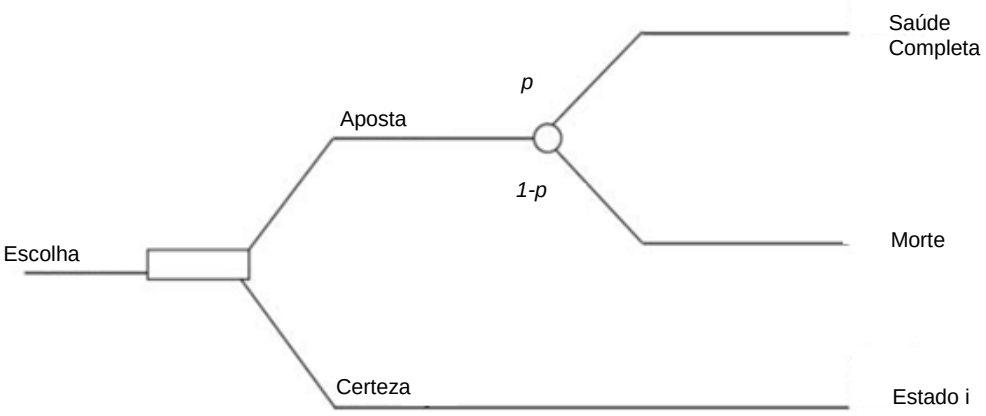

Figura 2. *Standard Gamble* para um estado de saúde crónico i, considerado melhor do que a morte (adaptado de [8]).

Na Figura 2 são representadas esquematicamente as duas opções caraterísticas do *Standard Gamble*. Uma das alternativas consiste num tratamento ou intervenção com dois resultados possíveis: ou o indivíduo regressa ao estado de saúde completa e vive por um período adicional de t anos, estando a este resultado associada uma probabilidade de p, ou morre imediatamente, situação que apresenta uma probabilidade de $1 - p$. [38] Esta alternativa tem, assim, uma incerteza associada. [38] A outra alternativa apresentada consiste na certeza de permanecer num determinado estado de saúde crónico conhecido, i, até ao fim da vida, isto é, durante t anos. [38]

Durante o processo de medição de utilidade recorrendo a este método, o indivíduo é questionado relativamente a quanto em termos de risco de morte está dispostos a aceitar para evitar a certeza do estado de saúde i, em avaliação. [54] Habitualmente esta questão não é colocada diretamente e segue-se um protocolo em que se vai variando a probabilidade de o indivíduo experienciar a morte, isto é a probabilidade p, de acordo com um processo iterativo tendo como base as respostas dadas, até o indivíduo indicar ser indiferente entre as duas alternativas apresentadas. [57]

O valor de preferência do estado de saúde i é, então, simplesmente igual à probabilidade de indiferença, ou seja, p^*, numa escala em que a utilidade do melhor estado (saúde perfeita) é 1 e a utilidade do pior estado (morte) é 0. [38] [57] Quanto mais severo for o estado de saúde i, maior será o risco de morte que o paciente está disposto a aceitar para conseguir atingir um estado de saúde completa. [8]

É possível aplicar este método a estados de saúde crónicos piores do que a morte e, também, a estados de saúde temporários, com algumas alterações ao formato anteriormente descrito. [15] [38]

Embora o *Standard Gamble* represente o método de medição de preferências com uma forte fundamentação teórica (de acordo com a teoria da utilidade) e seja um método extensivamente utilizado no apoio à tomada de decisão, apresenta como desvantagem o fato de as utilidades geradas deste modo serem dependentes do comportamento relativamente ao risco dos indivíduos inquiridos, sendo, mais concretamente, a utilidade uma função negativa do risco. [15] [52] [57] Geralmente, os indivíduos apresentam aversão ao risco, tendendo a sua escolha a recair sobre o resultado certo e conhecido ao invés de sobre a aposta incerta, produzindo, deste modo, valores de utilidade superiores aos obtidos comparativamente a indivíduos com um comportamento mais inclinado para seguir o risco. [56] Outra limitação associada a esta metodologia está relacionada com a dificuldade de entendimento e avaliação das probabilidades pelos indivíduos. Numa tentativa de colmatar estas dificuldades este método encontra-se, geralmente, associado a diferentes tipos de instrumentos visuais que pretendem ajudar na compreensão das probabilidades e, consequentemente, no desenvolvimento da tarefa de decisão. [15]

3.1.3. *Time Trade-Off*

O *Time Trade-Off* é um método de medição de preferências apenas concebido para aplicação na área da saúde e que apresenta como vantagem, relativamente ao *Standard Gamble*, o fato de não considerar situações de incerteza, não recorrendo assim a probabilidades e, como tal, não é influenciado pela atitude dos indivíduos perante o risco e ultrapassa as dificuldades sentidas por estes relativamente ao conceito de probabilidade. [54] Este método consiste, pois, numa tarefa de escolha entre dois cenários de saúde alternativos em que, ambos são caracterizados por resultados de saúde conhecidos e especificados. [38]

Para estados de saúde crónicos considerados melhor do que morte, a Figura 3 ilustra as duas alternativas apresentadas. A alternativa 1 consiste em viver t anos num estado de saúde debilitado, i, a ser avaliado, seguido por morte, enquanto a alternativa 2 consiste em viver com saúde completa por um menor período de tempo, x, seguido por morte. Os indivíduos são, assim, questionados sobre quantos anos de vida estão dispostos a sacrificar para evitar um determinado estado de saúde.

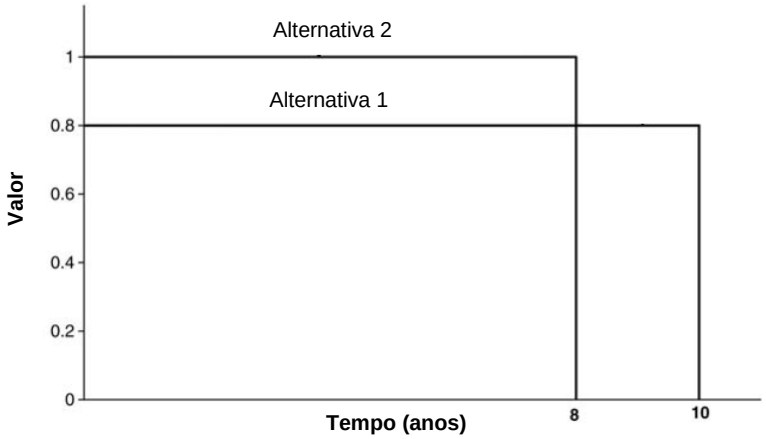

Figura 3. *Time Trade-Off* para um estado de saúde crónico, considerado melhor do que a morte (adaptado de [8]).

No protocolo de questionário TTO habitualmente seguido, o tempo vivido com saúde completa, x, vai sendo variado, de acordo com um processo iterativo tendo como base as respostas dadas, até que o indivíduo seja indiferente entre os dois cenários apresentados, sendo o tempo de indiferença x^*. [57] O valor de atribuído ao estado de saúde i é, assim, dado por x^*/t, numa escala em que o valor do melhor estado (saúde perfeita) é 1 e o valor do pior estado (morte) é 0. O valor atribuído ao estado de saúde i corresponde, assim, à fração de anos saudáveis equivalente a um ano passado nesse estado. [54]

Para estados de saúde crónicos piores do que morte e estados de saúde temporários é necessário proceder a alterações ao formato apresentado, continuando, no entanto, o seu princípio básico a recair sobre uma escolha entre duas alternativas conhecidas. [58] [59]

Conforme referido anteriormente, uma das principais vantagens associadas ao TTO consiste no fato de os valores gerados por este método serem indiferentes ao comportamento perante situações de risco, no entanto, os valores obtidos são suscetíveis à influência de preferências temporais, o que se apresenta como um problema, nomeadamente na escolha da duração do tempo de vida no estado de saúde a avaliar, t. [56]

Tal como, no SG também no TTO têm vindo a ser desenvolvidos instrumentos visuais de modo a facilitar a compreensão da tarefa pedida e reduzir as dificuldades sentidas pelos indivíduos.

3.1.4. Outros Métodos

3.1.4.1. *Person Trade-Off* (PTO)

A abordagem seguida no método *Person Trade-Off* difere de todas as abordagens até agora enunciadas, no sentido em que pretende estimar as preferências sociais dos diferentes estados de saúde e não as preferências individuais. [54] Tal como, no SG e no TTO é apresentada uma tarefa de escolha entre duas alternativas, no entanto, a diferença crucial reside no fato de que é pedida aos indivíduos uma tarefa de escolha num contexto de decisão envolvendo outros indivíduos e não o próprio. [57]

O protocolo de aplicação do PTO consiste em colocar ao indivíduo uma questão do tipo: ”Se existirem x pessoas numa situação de saúde A e y pessoas numa situação de saúde B, e se apenas se puder auxiliar (curar) um dos grupos, qual dos grupos escolheria ajudar?”. [38] Um dos valores x ou y deverá ser variado até que o indivíduo considere ambos os grupos indiferentes, podendo-se, nessa fase, calcular a indesejabilidade (ou desutilidade) da situação B como sendo x/y vezes maior do que a da situação A. [38]

Através de uma série de questões é possível relacionar as diferentes situações umas com as outras, de modo a obter uma escala de desutilidade que pode ser convertida numa escala intervalar de preferências, através da fórmula $y = 1 - x$. [38] [54]

3.1.4.2. *Ratio Scaling* (RS)

O método *Ratio Scaling* tira partido do fato de uma escala de desutilidade poder ser considerada uma escala de rácios com o estado saúde perfeita a representar o zero natural, ou seja, a não existência de desutilidade. [15] Assim, este método carateriza-se por permitir a construção de uma escala de desutilidade, através de comparações sucessivas entre pares de estados de saúde e indicação dos rácios de indesejabilidade (desutilidade) entre eles. [57]

Tal como no PTO a escala obtida pode, posteriormente, ser convertida numa escala intervalar de preferências. [38]

3.2. Medidas Genéricas Multi-Atributo Baseadas em Preferências

Embora indispensáveis na obtenção de índices de preferência associados à qualidade de vida relacionada com a saúde, os instrumentos diretos apresentados apresentam uma prática complexa, demorada e dispendiosa. [57] Como tal, nos últimos anos têm vindo a ser desenvolvidos instrumentos genéricos de avaliação de estados de saúde baseados na teoria Multi-atributo da Utilidade, que permitem a determinação de um valor de preferência associado ao estado de saúde de um indivíduo através de uma abordagem mais prática e simples. [38]

Os métodos indiretos de medição da QVRS baseada em preferências são constituídos por duas partes: um questionário, de auto-preenchimento, multi-atributo de classificação de estados de saúde que integra múltiplas dimensões de saúde e uma fórmula de pontuação baseada em preferências adequada, de modo a que a caracterização obtida recorrendo ao sistema de classificação multi-atributo possa ser combinada num único índice. [35]

Entre estes instrumentos destacam-se o *Quality of Well-Being* (QWB), o *Health Utilities Index* (HUI), o *EuroQol-5D* (EQ-5D) e o *Short-Form 6D* (SF-6D), sendo que, as principais diferenças entre eles recaem sobre o número de dimensões de saúde consideradas, o número e a descrição dos níveis em cada dimensão, a população na qual as preferências são baseadas e os instrumentos utilizados na sua medição. [8]

3.2.1. *Quality of Well-Being*

O *Quality of Well-Being* é um sistema genérico de classificação dos estados de saúde dos indivíduos de acordo com quatro atributos: mobilidade, actividade física, actividade social e sintoma-problema [15], em que cada atributo é descrito por 3 ou 5 níveis. [60] No caso de o indivíduo apresentar múltiplos sintomas ou problemas, aquele que for considerado mais indesejável deve ser o considerado. [15]

Os valores de utilidades dos estados de saúde são baseados em medições de uma amostra de população em geral recorrendo a uma escala de pontuação. Nesta tarefa de pontuação foi pedido aos indivíduos da amostra que avaliassem um dia nos diferentes estados de saúde e os colocassem numa escala, ancorada nos estados morte e saúde perfeita, de acordo com as suas preferências. [38]

Este instrumento foi o primeiro instrumento genérico baseado em preferências a ser desenvolvido [48] e tem vindo a ser aplicado em diferentes estudos e comparações com outros instrumentos. [15]

3.2.2. *Health Utilities Index*

As versões correntes do HUI são o HUI2 e o HUI3, que surgiram do desenvolvimento do HUI1. [60] Cada um destes instrumentos consiste num sistema de classificação de estados de saúde diferente, sendo definido por um conjunto de atributos, cada um com vários níveis variando entre não haver nenhuma incapacidade até haver incapacidade total, e uma fórmula de pontuação. [15] [48] Sendo que, tanto no HUI2 como no HUI3 as fórmulas de pontuação são baseadas em utilidades obtidas a partir da população em geral recorrendo ao método *Standard Gamble*. [61]

O HUI1 foi originalmente desenvolvido, baseando-se em parte no QWB, para avaliar resultado de saúde nos recém-nascidos com baixo peso. [60] Este instrumento é composto por 4 atributos: função física, incluindo mobilidade e actividade física; actividades habituais, incluindo cuidados pessoais e outras actividades habituais; função socio-emocional, incluindo estado emocional e actividades sociais; e problemas de saúde. [38] A valorização de utilidades foi obtida a partir da aplicação do *Time Trade-Off* e a fórmula de cálculo é uma função multi-atributo multiplicativa. [38]

O HUI2, por sua vez, foi desenvolvido para avaliar resultados a longo termo do cancro infantil. Este é um instrumento que tem em consideração 7 domínios: capacidade sensorial e de comunicação (compreende a visão, audição e fala); felicidade; cuidados pessoais, dor e desconforto; capacidade de aprendizagem e actividade física; e fertilidade. [38]

Por último, o HUI3 é, atualmente o mais utilizado, sendo constituído por 8 atributos cada um deles com 5 ou 6 níveis diferentes. [60] Os atributos neste instrumento são: visão; audição; fala; ambulatório; destreza; emoção; cognição e dor/desconforto. [38]

O HUI3 foi desenvolvido com base no HUI2, tendo o atributo fertilidade sido eliminado e o atributo da capacidade sensorial convertido em três atributos: visão, audição e fala. [15] Além disso, foram ainda efetuadas outras alterações com o objetivo de aumentar a independência estrutural dos atributos, isto é, se for concebível para um indivíduo situar-se num determinado nível de um atributo, independentemente dos níveis dos outros atributos. [38]

3.2.3. *EuroQol-5D*

O EQ-5D é o instrumento preferido de medição e valoração da qualidade de vida relacionada com a saúde nas avaliações efetuadas pelo NICE. [42] Este instrumento foi desenvolvido pelo Grupo EuroQol – um grupo multidisciplinar de investigadores – como um instrumento genérico de descrição e medição de qualidade de vida relacionada com a saúde sendo, habitualmente, utilizado em análises custo-utilidade. [62]

Este questionário permite a descrição de um problema ou condição de saúde de acordo com um sistema descritivo de cinco dimensões: mobilidade, autonomia, capacidade de desenvolver atividades

do dia-a-dia, dor e desconforto, ansiedade e depressão. Em que, para cada dimensão existem três estados possíveis, sendo eles, sem dificuldade, com alguma dificuldade ou com grandes dificuldades (Figura 21, Anexo A). [15]

Além da inclusão do sistema descritivo referido é, também, incluída no EQ-5D uma escala visual analógica, que permite aos indivíduos classificarem diretamente o seu estado de saúde de acordo com a sua perceção relativamente à sua qualidade de vida relacionada com a saúde. [62]

Através deste questionário é, assim, possível definir 243 (5^3) estados de saúde diferentes aos quais são adicionados os estados, "inconsciente" e "morto". [38] Para determinação do algoritmo de pontuação adequado a este instrumento, um subconjunto de estados de saúde foi avaliado por uma amostra da população em geral recorrendo ao método *Time Trade-Off*, tendo sido posteriormente estimados modelos para obter valores para todos os estados de saúde que podem ser definidos de acordo com esta metodologia. [15]

A utilização do questionário EQ-5D é cognitivamente simples e requer um curto espaço de tempo, podendo este ser aplicado por autopreenchimento via correios, em ambiente clinico ou em entrevistas cara-a-cara. [8]

3.2.4. *Short-Form 6D*

O SF-6D é um instrumento de medição da qualidade de vida relacionada com a saúde baseado em preferências desenvolvido recentemente e que tem vindo a ser cada vez mais a ser utilizados em diferentes estudos. [63]

Este instrumento, tal como os anteriormente referidos, é composto por um sistema classificação multi-atributo de estados de saúde e por uma função de pontuação. Apresentado a particularidade de que o sistema de classificação multi-atributo tem como base o popular questionário de avaliação de qualidade de vida relacionada com a saúde, SF-36, e também no SF-12. [15] O sistema de classificação é, assim, composto por seis dimensões (função física, limitação no desempenho, função social, dor física, saúde mental e vitalidade) obtidas recorrendo a onze itens do SF-36 ou a oito itens do SF-12. [15] Cada dimensão apresenta quatro a seis níveis, dando origem a um total de 18 000 estados de saúde diferentes. [15]

O algoritmo de pontuação associado a este instrumento foi desenvolvido com base no método *Standard Gamble*, aplicado a uma amostra geral da população, obtendo-se, assim, diretamente valores de utilidade para alguns dos 18 000 estados de saúde possíveis, tendo, posteriormente sido estimados modelos econométricos para atribuir peso aos diferentes níveis das dimensões do SF-6D, permitindo a obtenção de valores para todos os estados definidos por este instrumento [35]

3.3. Outros Métodos de Obtenção de Utilidades em Saúde

Como referido anteriormente, o interesse crescente relativamente à inclusão dos pacientes e da população em geral na tomada de decisões em saúde levou ao desenvolvimento de diversos métodos de elicitação de preferências em relação a diferentes estados de saúde. [45]

Estes métodos podem ser classificados em métodos cardinais e métodos ordinais. [64] Os métodos de medição de utilidades até agora analisados (VAS, SG e TTO) são classificados como métodos cardinais, uma vez que, geram preferências de forma quantitativa baseadas nas escolhas feitas pelos indivíduos inquiridos. [64] [65] Estes métodos permitem, assim, obter uma estimativa numérica direta das utilidades, ou seja, do grau com que um determinado estado de saúde é preferível a outro. [64] Em economia de saúde, estes são os métodos habitualmente utilizados na medição de utilidades, sobretudo na validação de instrumentos genéricos multi-atributo baseados em preferências.

Os métodos ordinais de elicitação de preferências, por sua vez, estão interessados na ordenação de preferências entre duas ou mais alternativas sem estabelecer diretamente o grau de preferência de umas sobre as outras. [64] Destacam-se, nesta categoria, os *Discrete Choice Experiments*, uma metodologia que apresenta um rápido crescimento em termos de literatura na área da saúde, nos últimos anos. [66]

Os métodos de *Discrete Choice Experiments* podem definir-se como 'uma abordagem geral de elicitação de preferências em que a tarefa pedida aos indivíduos inquiridos recai sobre uma escolha ou escolhas, entre duas ou mais alternativas discretas, em que pelo menos um atributo da alternativa é sistematicamente variado entre indivíduos inquiridos de tal forma que a informação relacionada com os parâmetros de preferência de uma função de utilidade indireta pode ser inferida'. [65] Estes métodos permitem, assim, a transformação de medidas ordinais em medidas cardinais úteis, por exemplo, nas avaliações económicas em saúde. [64]

Tipicamente os DCEs envolvem, assim, um questionário onde é pedido aos indivíduos que escolham entre opções, hipotéticas ou reais, descritas usualmente por ou mais atributos, sendo que, de acordo com a classificação estabelecida em [65], distinguem-se dentro desta abordagem métodos conceptualmente diferentes, com base no formato da questão colocada aos indivíduos.

Quando a tarefa pedida consiste numa simples escolha binária para uma alternativa estamos perante o *Binary Choice Experiment*. [45] Sendo este o método mais utilizado atualmente. [45] Por outro lado, quando o número de alternativas em escolha é superior a 3, isto é, temos um *Multinomial Choice Experiment.* [65]

Outro método que, segundo alguns autores tem uma história separada, dos DCEs é o *Ranking exercise* [64], que como o nome indica consiste em colocar por ordem um conjunto de alternativas ou atributos, baseado nas preferências dos indivíduos. [65] É defendido em [65] que este método seja considerado um DCEs, uma vez que, a definição de um ranking de preferências reflete a comparação entre todos os pares de alternativas possíveis.

Por último, destaca-se o *Best-Worst Experiment,* um método que embora permita obter menos informações do que o *Ranking* envolve menos dificuldades e permite obter mais informações do que o BCE. Neste método é pedido aos indivíduos inquiridos que identifiquem a melhor e a pior alternativa do conjunto, isto é, aquelas que apresentam maior e menor preferências, respetivamente. [65]

Os DCEs apresentam, então, como principal vantagem relativamente aos métodos de medição de preferências cardinais o fato de serem de mais fácil compreensão, exigindo um menor esforço

cognitivo aos indivíduos e, como tal, apresentam uma administração facilitada, além de apresentarem maior confiança derivada de erros de medição reduzidos. [64]

Embora os DCEs possam ser utilizados na medição de preferências relativamente a estados de saúde, atualmente, a sua maior utilização é relativa a avaliação se serviços de saúde, produtos, práticas, intervenções e políticas. [64] [67] Além disso, os atributos das diferentes alternativas podem ser processos ou resultados e podem ou não estar relacionados com a saúde, tornando possível desse modo avaliar outros benefícios além dos diretamente refletidos na saúde obtida. [68]

Além destes métodos, os métodos de decisão multicritério (MCDM) têm vindo, também, a merecer destaque na área da saúde. [3] [69] Os MCDM fornecem aos decisores uma abordagem de apoio à decisão estruturada e transparente que permite identificar a alternativa com maior preferência, tendo em conta múltiplos critérios de análise e a sua importância relativa, bem como, a *performance* das alternativas nesses critérios. [70]

Os MCDM podem ser classificados em diferentes categorias sendo, no contexto deste trabalho, destacados os modelos de medição de valores. [69] Este método permite considerar explicitamente as preferências do público ou dos pacientes na tomada de decisões relativas a cuidados de saúde [45], sendo o grau com que uma alternativa é preferida sobre outra representado através da construção e comparação de valores numéricos. [69] Diferentes valores numéricos são desenvolvidos, separadamente, para os vários (pares de) critérios em análise e posteriormente são agregados num único valor. [45] Estes métodos distinguem-se assim dos anteriores pelo fato de as preferências serem determinadas tendo em conta diferentes critérios e não ser pedido aos inquiridos que atribuam valor a uma alternativa como um todo. [45]

Na elicitação de preferências em saúde distinguem-se dois processos diferentes, os processos baseados na teoria de utilidade multi-atributo e os *Analytic Hierarchy Process* (AHP). [45]

Os métodos baseados na teoria de utilidade multi-atributo, uma extensão da teoria da utilidade esperada, determinam as preferências de uma alternativa através da medição de um conjunto de critérios específicos e do mapeamento dos resultados obtidos numa escala de utilidades através de uma função de pontuação específica. [61] [71] A ordem de importância e os pesos dos atributos são determinados através da apresentação aos decisores de *trade-offs* explícitos entre apostas. [45] Entre estes encontram-se os instrumentos genéricos de medição baseados em preferência, anteriormente referidos, que permitem avaliar a utilidade de diferentes estados de saúde.

O *Analytic Hierarchy Process* consiste noutro processo de MCDM em que os pesos relativos de cada critério não são atribuídos diretamente mas através de comparações entre pares de critérios, através das quais uma escala de valores que mantem a informação obtida é estimada. [45] Contudo, este processo apresenta diferentes críticas, nomeadamente em relação ao método de valor próprio usado na derivação das prioridades. [72]

3.4. Estudos comparativos de métodos diretos de obtenção de utilidades

A avaliação de preferências em saúde é, atualmente, uma área em desenvolvimento, tendo vindo a surgir não só novas metodologias mas também diferentes estudos de exploração e comparação entre metodologias já existentes, com o objetivo de analisar os diferentes fatores que fazem variar sistematicamente as preferências obtidas, nomeadamente, o método a que se recorre, a população à qual é aplicado, o modo de apresentação e descrição dos estados de saúde a avaliar e o contexto em que são enquadradas as questões. [30]

Na Tabela 2 é feita referência a alguns estudos comparativos entre diferentes métodos de avaliação de preferências em saúde, que sustentam a discussão relativa a qual o método mais adequado a utilizar em diferentes situações.

Tabela 2. Estudos comparativos entre diferentes métodos diretos de obtenção de utilidades em saúde.

Estudo	Objetivo e Métodos
[52] "*Valuing Health States: A comparison of methods"*	Comparação direta entre TTO e SG
[58] "*A comparison of three valuation methods for temporary health states in patients treated with oral anticoagulants*"	Comparação entre o método convencional TTO, uma adaptação do mesmo para estados de saúde temporários e uma escala de pontuações
[43] "*Measuring Preferences for Health States Worse than Death*"	Comparação entre resultados obtidos para estados de saúde considerados piores do que a morte recorrendo ao SG, TTO e Escala Categórica
[73] "*A comparison of directly elicited and pre-scored preference-based measures of quality of life: the case of adhesive capsulitis*"	Comparação entre o método direto TTO e métodos indiretos de obtenção de índices de QVRS
[74] "*The Comparability and Reliability of Five Health-State Valuation Methods*"	Comparação entre valores e utilidades obtidas recorrendo ao SG, TTO, Escala de pontuações e WTP (*Willingness-to-pay*)

Além da discussão relativamente ao método de medição de preferências a utilizar, é também importante ter em consideração a amostra da população sobre a qual se vão considerar as preferências, existindo evidências de que as utilidades obtidas variam consoante o grupo populacional que as avalia. [9] [40]

São dois os grupos de indivíduos cujas preferências são habitualmente relevantes, os pacientes e a população em geral. [48] Geralmente, quando se pede aos pacientes para avaliarem o seu próprio estado de saúde as utilidades obtidas são superiores às utilidades elicitadas pela população em geral, para os mesmos estados de saúde. [75] Verifica-se, também, quando é pedido aos pacientes que avaliem estados hipotéticos comparativamente piores que seu estado atual, que as utilidades obtidas tendem a ser superiores às elicitadas pela população em geral. [56]

Alguns estudos realizados nesta área revelam que as discrepâncias detetadas estão sobretudo associadas à forma como os diferentes estados de saúde são descritos no caso de avaliação pela população e à forma como diferentes indivíduos percebem e interpretam a informação, ou seja, como a experiência vivida anteriormente ou durante o processo de avaliação, ou até mesmo as crenças pessoais e estereótipos, levam a diferentes ajustamentos. [9]

Assim, conclui-se que embora sejam os pacientes o grupo com maior capacidade de avaliar o seu próprio estado de saúde quando se trata da avaliação de estados hipotéticos esta é influenciada pelos estados de saúde atuais. Por outro lado, a avaliação pela população em geral apesar de apresentar limitações relativamente ao entendimento dos diferentes estados de saúde e ser também influenciada pelo estado saudável de saúde atual deve ser tida em conta, sobretudo, num sistema de saúde público onde são os recursos da população que estão a ser alocados. [56] Além disso, membros da população em geral são potenciais pacientes do futuro para a maioria dos programas ou dos resultados de saúde em estudo nestas análises. [55]

A definição da amostra a estudar deve, então, ser feita com base no contexto de tomada de decisão e no propósito da análise. Sendo, atualmente, no caso de avaliações económicas o grupo geralmente utilizado o da população em geral uma vez que estas análises pretendem orientar políticas sociais e não as decisões individuais dos pacientes [41] [75]

3.5. Medição de QVRS em Portugal

Uma vez que o trabalho desenvolvido nesta tese pretende desenvolver um novo método de avaliação de estados de saúde e aplicá-lo ao contexto português, considerou-se relevante a realização de uma pesquisa sobre os desenvolvimentos da área em Portugal, cujos resultados principais são apresentados de seguida.

Em Portugal, a avaliação de tecnologias de saúde tem acompanhado a tendência de crescimento e desenvolvimento verificada em vários outros países, nomeadamente, nos países da União Europeia. Dentro da ATS, a avaliação económica tem adquirido uma importância crescente, nomeadamente as análises custo-utilidade com recurso ao uso de QALYs como medida de ganhos em saúde, que têm vindo a ser cada vez mais utilizadas na informação da tomada de decisão em saúde. [76]

Relativamente à abordagem dos QALYs um dos assuntos de maior interesse para os investigadores de economia de saúde é a medição da qualidade de vida relacionada com a saúde, nomeadamente, as metodologias que permite captar as preferências dos indivíduos relativamente a diferentes resultados de saúde e a incorporação dos resultados obtidos nas análises efetuadas. Como já foi referido anteriormente, nos últimos anos têm sido desenvolvidos vários instrumentos genéricos de medição de preferências, baseados na teoria multi-atributo de utilidade, que permitem calcular, de forma simples, os valores de utilidade que os indivíduos atribuem a diferentes estados de saúde. Destacando-se, entre estes, o EQ-5D, o SF-6D e o HUI.

Em relação a este tipo de instrumentos, existe um interesse crescente no estudo da variação, entre países, de valores de utilidades para diferentes estados de saúde obtidos, tendo aumentado o

número de estudos realizados com o propósito de definir sistemas de valores próprios bem como normas populacionais. [77]

A definição de sistemas de valores próprios para cada país e de normas populacionais é uma necessidade para os investigadores nesta área, proporcionando não só o acesso a um sistema de valores adaptado às especificidades e cultura da população em estudo mas também normas populacionais específicas bastante úteis na contextualização e interpretação de diferentes resultados de investigação obtidos recorrendo a estes instrumentos, servindo como valores de referência.

Em Portugal, existe atualmente um sistema de valores e normas populacionais definidos para dois dos instrumentos de medição de preferências mais utilizados, o EQ-5D e o SF-6D.

A versão portuguesa do questionário EQ-5D foi criada em 1997, pelo Centro de Estudos e Investigação em Saúde da Universidade de Coimbra, com base em normas de orientação definidas pelo Grupo EuroQol, incluindo os procedimentos de tradução e retroversão. [78] Posteriormente, esta versão foi alvo de alterações num processo de harmonização entre as várias versões em língua de origem latina e apenas em 1998 foi disponibilizada a versão portuguesa final do EQ-5D. [78]

Após a sua finalização este instrumento foi amplamente utilizado, no entanto, apenas no ano de 2013 foram publicados os primeiros estudos que conduziram inicialmente à versão portuguesa e à garantia de aceitabilidade, fiabilidade e validade. [78] Foram também publicados nesse ano os resultados obtidos de um grande projeto de medição e valoração da saúde através do EQ-5D em Portugal. Estes resultados foram publicados em dois estudos, [76] e [79], tendo o primeiro estudo publicado como objetivo validar o EQ-5D para a realidade portuguesa, determinando um sistema de valores português através de uma avaliação de diferentes estados de saúde, por parte de uma amostra da população em geral, recorrendo ao *Time Trade-Off*. Foi, ainda, comparado o sistema de valores obtido com os de outros países, tais como, Reino Unido, França, Espanha e Estados Unidos da América tendo-se concluído que o sistema de valores português difere de todos estes sendo que, o mais próximo é o sistema de valores espanhol ao invés do sistema de valores do Reino Unido, que na ausência do português foi o utilizado.

O segundo estudo publicado teve como objetivo apresentar os valores obtidos de qualidade de vida relacionada com a saúde da população portuguesa, recorrendo ao EQ-5D, ou seja, pretendeu apresentar as normas populacionais portuguesas para este instrumento. O valor médio de utilidade obtido para a população geral foi de 0.758, tendo-se registado variações entre -0.50 e 1.00. Foram também calculadas as normas populacionais de acordo com o género, grupo etário, estado civil, região, nível de escolaridade, estado ocupacional, dimensão do agregado familiar e existência ou não de doenças crónicas.

Em relação à versão portuguesa do SF-6D esta foi publicada em 2006, num estudo onde se investigaram pela primeira vez as preferências sobre estados de saúde da população portuguesa. [80] Em 2009 foi publicado o primeiro sistema de valores português do SF-6D que veio em substituição da utilização do sistema de valores britânico [81], no entanto, esse sistema apresentava algumas limitações, nomeadamente, a existência de incoerências nos pesos de alguns níveis de dimensão, tendo-se em 2011 publicado um novo estudo ([77]) que permitiu obter um sistema de valores melhorado. A obtenção do sistema de valores para o SF-6D é feita recorrendo ao *Standard*

Gamble, aplicado a uma amostra da população em geral, para avaliar as preferências da população relativamente a um conjunto de estados de saúde. [77] Neste último estudo foram, ainda, definidas as normas populacionais portuguesas para o SF-6D, tendo-se obtido um valor de utilidade média dos estados de saúde da população ativa portuguesa, dos 18 aos 64 anos, de 0.81, associada a um desvio padrão de 0.12. Foram também calculadas as normas populacionais de acordo com o género, grupo etário, estado civil, região, nível habitacional, nível de escolaridade e existência ou não de doenças crónicas.

Após a definição, para instrumentos como o EQ-5D e SF-6D, de um sistema de valores e de normas populacionais para Portugal, baseados numa amostra da população em geral, segundo alguns autores é necessária a validação dos sistemas de valores obtidos em amostras de doentes e em subgrupos culturais existentes na população em geral. [77] Alguns estudos deste tipo têm vindo a ser desenvolvidos em Portugal, destacando-se a validação em amostras de indivíduos com doenças com elevada prevalência, que se apresenta na Tabela 3.

Tabela 3. Estudos de medição de QVRS baseada em preferência realizados em Portugal.

Estudo	Doença em causa	Instrumento de Medição de Preferências Utilizado
[82] [83]	Artrite Reumatóide	SF-6D EQ-5D
[49] [83]	Cataratas	EQ-5D
[84] [83]	Asma	SF-6D EQ-5D
[83]	Doença pulmonar Obstrutiva Crónica	EQ-5D
[85]	VIH	EQ-5D

3.6. Influência da Numeracia e da Fluência na Expressão de Julgamentos de Valor

O crescente desenvolvimento de diferentes metodologias de elicitação de preferências em diferentes áreas da tomada de decisão e a possibilidade de escolha entre o recurso a metodologias recorrendo a julgamentos de valor numéricos ou julgamentos de valor não numéricos levantou questões relativamente à influência que competências como a numeracia e a fluência verbal dos indivíduos podem ter, não só, na forma como estes preferem expressar as suas preferências mas também nos próprios resultados obtidos. [12]

Embora as diferenças de preferências entre expressões de valor não tenham sido ainda amplamente abordadas pela literatura, nomeadamente na área da saúde, existem alguns estudos que revelam que a forma como a informação numérica e verbal acerca de diferentes produtos é apresentada influencia a informação processada pelos indivíduos. [86] Outros estudos revelam, ainda, que o formato que os decisores preferem usar para expressar probabilidades e para as

receber também é influenciado pela sua competência numérica. [12] Nomeadamente na área da saúde foi, também, estudado o impacto da numeracia na capacidade dos pacientes compreenderem situações de risco e benefício caraterizadas numericamente, tendo-se determinado que uma maior numeracia está fortemente relacionada com uma interpretação e um uso mais correto de informação quantitativa apresentada. [87]

Num estudo recente apresentado em [12] , pretendeu-se examinar a influência da numeracia e da fluência na perceção e preferências dos decisores relativamente a duas técnicas de elicitação de valores, uma numérica e outra não numérica. O resultado obtido esteve de acordo com o esperado, ou seja, que quanto maior a numeracia maior a preferência por métodos numéricos e, por outro lado, quanto maior a fluência verbal maior a preferência pela realização de julgamentos com palavras.

Adicionalmente, alguns estudos foram já desenvolvidos com o intuito de avaliar a influência da numeracia nos métodos convencionais de elicitação de preferências em saúde, todos eles baseados em julgamentos de valor numéricos. [10] [11] A partir destes concluiu-se que, existe uma maior correlação entre diferentes métodos de elicitação de preferências para indivíduos numerados, sendo esta correlação muito inferior para amostras populacionais não numeradas.

Tendo em conta os principais objetivos definidos neste trabalho, no presente capítulo foi apresentada uma revisão bibliográfica dos métodos de medição de preferências em saúde comumente encontrados na literatura, destacando-se algumas das suas limitações, bem como, alguns novos métodos que têm sido desenvolvidos. Foi feita, ainda, referência aos desenvolvimentos na área em Portugal e a estudos recentes que comprovam a influência de competências dos indivíduos, tais como, a numeracia e a fluência, na expressão de julgamentos de valor, sendo este último aspeto um ponto de relevo na motivação para o desenvolvimento do presente estudo.

Uma vez concluída a contextualização do presente trabalho e a apresentação e discussão dos métodos de avaliação de estados de saúde atualmente existentes, dá-se início, no capítulo seguinte, à descrição da metodologia proposta neste estudo.

.

4. Metodologia Proposta

Na revisão bibliográfica, apresentada no capítulo anterior, foram identificados e descritos os principais métodos de avaliação da QVRS baseados em preferências, comumente encontrados na literatura. Relativamente aos métodos diretos de valoração de estados de saúde evidenciaram-se algumas das suas limitações teóricas e empíricas, nomeadamente algumas questões que têm sido levantadas no que diz respeito à sua complexidade de administração e elevada exigência cognitiva, o que tem conduzido ao desenvolvimento de novos métodos baseados, sobretudo, em julgamentos de valor não numéricos. Adicionalmente referiram-se alguns estudos recentes que apontam para o fato de competências, como a numeracia e a fluência, terem impacto no modo preferencial de expressar julgamentos de valor (em números ou palavras) e nos próprios julgamentos de valor obtidos, nomeadamente estudos na área da saúde que determinam a influência da numeracia nas utilidades obtidas recorrendo a métodos convencionais, baseados na atribuição de julgamentos de valor numéricos.

Neste contexto, afigura-se como relevante desenvolver e testar um novo método de avaliação de estados de saúde baseado em julgamentos não numéricos, bem como, proceder à comparação entre os resultados obtidos com este novo método e com um método convencional baseado em julgamentos numéricos. Paralelamente é de interesse investigar as seguintes hipóteses relacionadas com a numeracia:

1) Existe uma relação entre o modo preferencial de expressar julgamentos de valor de um indivíduo, em números ou palavras, e o seu grau de numeracia? De acordo com Fasolo e Bana e Costa [12] a expressão de julgamentos de valor em números ou palavras não é uma tarefa psicologicamente equivalente para indivíduos com diferentes graus de numeracia e de fluência. Indivíduos com maior grau de numeracia têm preferência por métodos de valoração que recorram à atribuição de julgamentos de valor numéricos, enquanto indivíduo com maior grau de literacia têm maior preferência por métodos baseados na atribuição de julgamentos de valor não numéricos.
2) O grau de numeracia dos indivíduos tem influência nas utilidades obtidas? De acordo com Woloshin *et al.* [10] e Schwartz *et al.* [11] as utilidades médias obtidas diferem de acordo com o grau de numeracia da amostra populacional verificando-se que, enquanto, os resultados obtidos para a amostra com elevado grau de numeracia estão de acordo com o esperado, no contexto do estudo em questão, o mesmo não se verifica para os resultados obtidos pela amostra com baixo grau de numeracia.
3) As escalas de valores obtidas por diferentes métodos apresentam maior correlação para amostras populacionais com maior grau de numeracia? Ainda de acordo com Schwartz *et al.* [11] determinou-se uma maior correlação entre utilidades obtidos por métodos diferentes para indivíduos com elevada numeracia do que para indivíduos com baixa numeracia.

Este capítulo começa por apresentar uma visão geral da metodologia desenhada para responder aos objetivos definidos. As secções seguintes explicam com detalhe cada componente da metodologia.

4.1. Desenho Metodológico

A metodologia proposta de modo a atingir os objetivos anteriormente definidos nesta tese e dar resposta às hipóteses levantadas relativamente à numeracia é, em síntese, apresentada na Figura 4, consistindo num conjunto de etapas progressivas que fazem uso de diferentes ferramentas.

A etapa inicial, que diz respeito à fase conceptual do estudo, consistiu numa recolha de informações com o objetivo de caraterizar o contexto em que o estudo se insere, definir as principais motivações e tomar conhecimento de literatura relevante na área, nomeadamente, relativa a diferentes métodos de avaliação de preferências em saúde baseados em julgamentos de valor numéricos e não numéricos. Simultaneamente à recolha de informações procedeu-se à formulação do estudo a ser desenvolvido e à identificação dos seus principais objetivos.

A etapa seguinte dá início à metodologia proposta propriamente dita e, consiste no desenho de um novo protocolo de avaliação de estados de saúde com base em julgamentos qualitativos recorrendo-se, para tal, à abordagem MACBETH. Esta abordagem carateriza-se por permitir a obtenção de uma escala cardinal de preferências (ou valor relativo) tendo como base a atribuição de julgamentos semânticos de diferença de atratividade entre pares de opções. [88] A abordagem MACBETH, bem como o novo protocolo proposto para avaliação de estados de saúde são descritos em detalhe na Secção 4.2.

A principal motivação para a utilização desta abordagem na valoração de estados de saúde consiste no fato de esta refletir um processo cognitivamente mais simples do que aquele requerido pelos métodos convencionais, que não implica a atribuição de julgamentos de valor numéricos e, como tal, não é susceptível à influência da numeracia da população. Além disso, a administração do novo protocolo proposto é também simplificada, uma vez que não recorre a protocolos distintos para estados melhores e piores do que a morte e não recorre a processos iterativos demorados para identificação do ponto de indiferença entre duas alternativas, característicos de métodos como o TTO e o SG. Este método constitui, assim, uma alternativa aos métodos convencionais.

Com o objetivo de implementar o novo protocolo de avaliação de estados de saúde proposto, bem como, o método numérico com a qual se pretende fazer a comparação dos resultados obtidos, propõe-se o desenvolvimento de um questionário *online*. O desenvolvimento deste questionário consistiu numa das principais etapas da metodologia proposta, envolvendo várias atividades, cuja ordem não é unidirecional sendo possível retroceder e introduzir ajustamentos sempre que necessário:

- Definição do método convencional de elicitação de preferências, baseado em julgamentos numéricos, a implementar conjuntamente com a nova metodologia proposta;
- Definição do sistema de caraterização a utilizar na descrição dos diferentes estados de saúde a avaliar e seleção dos mesmos;
- Estruturação do questionário e definição dos diferentes tipos de informação que se pretende recolher: caraterização do estado de saúde atual dos indivíduos; recolha de julgamentos numéricos e não numéricos para avaliação de estados de saúde hipotéticos; avaliação do

grau de numeracia; determinação da preferência entre os dois tipos de julgamentos implementados e caraterização socioeconómica;

- Escolha da plataforma Web para implementação do questionário e construção do mesmo tendo em conta as ferramentas disponíveis e questões metodológicas de aleatorização, nomeadamente relativas à ordem de apresentação das tarefas de atribuição de julgamentos numéricos e não numéricos e relativas aos estados de saúde apresentados para avaliação (conforme descrito nas Secções 4.3.3 e 4.3.2, respetivamente);
- Definição da estratégia de distribuição e da população-alvo;

Após a implementação do questionário na plataforma selecionada, a etapa seguinte consistiu, em primeiro, na sua distribuição de forma limitada a uma amostra de teste, com o intuito de verificar o correto funcionamento de todas as suas funcionalidades e receber críticas construtivas, e, em segundo, no melhoramento do questionário de acordo com as sugestões recebidas. Uma vez obtido o questionário na sua versão final prosseguiu-se com a sua distribuição, de acordo com a estratégia de distribuição e a população-alvo anteriormente definidas.

Concluída a recolha de dados, a etapa seguinte consistiu na valoração dos estados de saúde propriamente dita, tendo em conta os julgamentos numéricos e não numéricos recolhidos, de acordo com cada um dos métodos implementados. Simultaneamente com esta etapa foram, também, identificados e aplicados critérios de exclusão aos questionários recolhidos.

A última etapa da metodologia proposta consistiu na análise dos resultados obtidos, de modo a dar resposta a diferentes pontos:

- Descrição do processo de obtenção da amostra populacional válida para análise, incluindo a caraterização dos motivos de exclusão e análise da influência da ordem de apresentação das tarefas de atribuição de julgamentos numéricos e não numéricos;
- Caraterização da amostra relativamente às variáveis sociodemográficas, ao estado de saúde atual e a grau de numeracia;
- Análise comparativa dos resultados obtidos recorrendo aos dois métodos de avaliação de estados de saúde implementados;
- Caraterização das preferências entre os dois tipos de julgamentos implementados;
- Influência da numeracia: nas preferências manifestadas, nas escalas de valores obtidas e nas exclusões efetuadas.

As várias etapas necessárias ao desenvolvimento da metodologia proposta encontram-se descritas, com maior detalhe, nas secções seguintes.

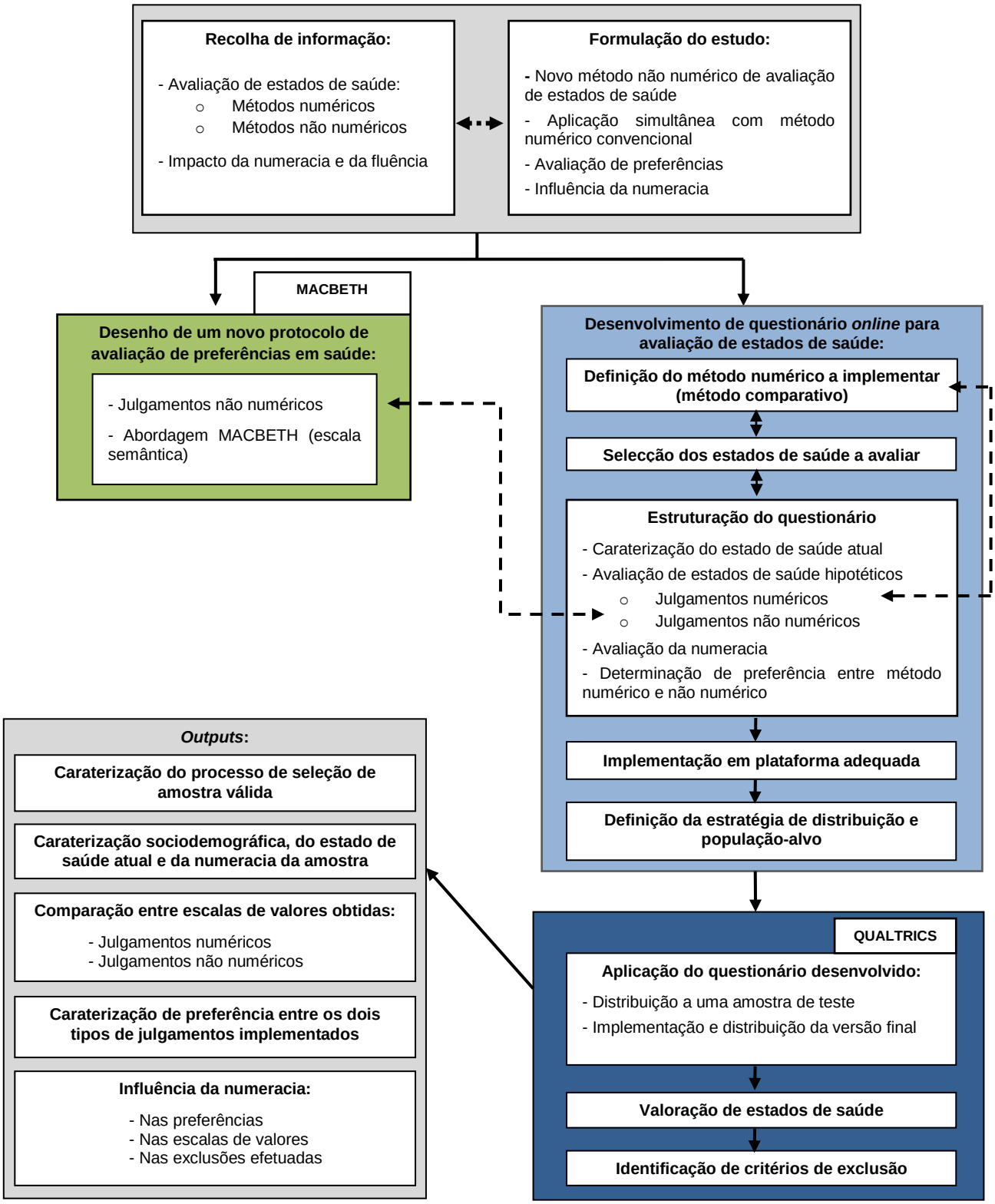

Figura 4. Representação esquemática da metodologia proposta. As setas contínuas representam relações sequenciais e as setas a tracejado representam processos iterativos.

4.2. Abordagem MACBETH

O MACBETH (***M**easuring **A**ttractiveness by a **C**ategorical **B**ased **E**valuation **T**echnique*) é uma abordagem não numérica, amplamente utilizada no apoio à tomada de decisão em diferentes áreas, que permite avaliar diferentes opções tendo em conta múltiplos critérios baseando-se, tal como o nome indica, na medição de atratividade por uma técnica de avaliação baseada em categorias para ponderar critérios e gerar pontuações para as opções em cada critério. [88] [89] Esta abordagem permite, assim, a obtenção de escalas cardinais de atratividade (ou valor relativo) das opções através da recolha de julgamentos qualitativos relativos a diferenças de atratividade.

De acordo com a abordagem MACBETH, na obtenção de julgamentos, os elementos de um dado conjunto em análise são apresentados dois a dois, com o primeiro melhor do que o segundo, e é pedido aos indivíduos que indiquem a diferença de atratividade entre eles recorrendo a uma escala semântica constituída por sete categorias: *nula*, *muito fraca*, *fraca*, *moderada*, *forte*, *muito forte* ou *extrema*. [88] [90] Tendo em conta os julgamentos qualitativos de diferença de atratividade obtidos, o MACBETH propõe uma escala numérica de valores para as opções, através da resolução de um problema de programação linear. [88]

Esta abordagem permite, então, obter através da comparação de opções duas a duas em termos qualitativos, informação cardinal em relação ao valor relativo de cada opção, pelo que a sua adaptação ao contexto deste estudo se revela adequada tendo em conta as principais motivações anteriormente definidas. A abordagem MACBETH, habitualmente usada em modelos multicritério de avaliação foi, assim, adaptada como método de avaliação de estados de saúde seguindo-se, neste caso, uma abordagem monocritério em que o conjunto de opções em análise diz respeito aos diferentes estados de saúde caraterizados em múltiplas dimensões. Nas subsecções seguintes o processo técnico de elicitação de valor das opções é descrito com maior detalhe (Secção 4.2.1.), bem como, a adaptação da abordagem MACBETH como método de avaliação de estados de saúde a ser desenvolvido e testado neste estudo (Secção 4.2.3.).

4.2.1. Questionamento MACBETH e obtenção de Escalas de Valor

Conforme referido anteriormente, a abordagem MACBETH pode ser usada para construir uma escala de valores partindo de uma escala categórica, através de um processo de comparações entre pares de elementos requerendo julgamentos qualitativos relativos a diferenças de preferências. [70] Assim, considerando-se duas opções, em que a opção x é mais atrativa do que a opção y, pretende-se obter um julgamento qualitativo, relativo à diferença de atratividade entre elas, tendo por base uma escala semântica constituída por sete categorias: *nula* (quando não há nenhuma diferença de atratividade entre as opções), *muito fraca*, *fraca*, *moderada*, *forte*, *muito forte* ou *extrema*. [88] Em casos de hesitação ou divergência é possível a seleção de várias categorias consecutivas, exceto para a categoria *nula*. [88]

Os julgamentos obtidos para os pares ordenados de opções (x, y) vão sendo introduzidos no *software* M-MACBETH [89], de modo a preencher uma matriz designada matriz de julgamentos semelhante à esquematizada na Figura 5.

(a)

	a=sup.	b	c	d	e=inf.
a=sup.	nula	mt. fraca	moderada	positiva	positiva
b		nula	frac-mod	moderada	positiva
c			nula	fraca	positiva
d				nula	nula
e=inf.				nula	nula

extrema
mt. forte
forte
moderada
fraca
mt. fraca
nula

(b)

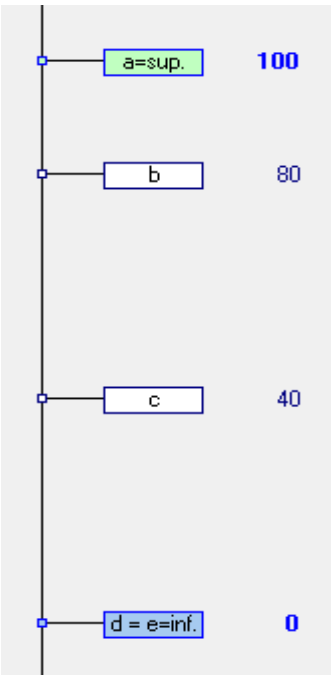

Figura 5. (a) Exemplo de matriz de julgamentos MACBETH consistente; **(b)** Escala proposta pelo *software* M-MACBETH (adaptado de [88]).

O preenchimento da matriz de julgamentos não requer nenhuma ordem específica nem a atribuição de julgamentos a todos os pares possíveis de opções. Considerando um conjunto X com n opções, se estas estiverem ordenadas por preferência decrescente, a obtenção de uma escala de valores válida é conseguida através do preenchimento do triângulo superior da matriz, o que corresponde a um número máximo de $n(n-1)/2$ julgamentos. No entanto, o número mínimo teoricamente aceitável de julgamentos é $(n-1)$, correspondendo estes à comparação entre opções ordenadas consecutivamente ou à comparação de uma só opção com todas as outras. [90] Porém, na prática, é aconselhável que sejam introduzidos julgamentos adicionais, de modo a que seja possível testar a consistência dos julgamentos introduzidos na matriz. [88]

Relativamente às entradas na matriz para as quais existe apenas informação ordinal e nenhum julgamento de diferença de atratividade foi atribuído ao par de opções, o M-MACBETH atribui como julgamento uma diferença de atratividade "*positiva*", o que significa que para este par o tipo de avaliação disponível é apenas ordinal. [88]

À medida que os julgamentos qualitativos são obtidos e introduzidos na matriz de julgamentos, o *software* testa automaticamente a sua consistência com os restantes julgamentos já introduzidos, e sempre que é encontrada uma inconsistência identifica a origem do problema e oferece sugestões para eliminá-lo tendo em conta o menor número de modificações necessárias. [90]

Para que uma matriz de julgamentos seja consistente deve ser possível, a partir dos julgamentos introduzidos, obter uma escala de pontuações incluindo todos as opções em análise, de modo que: [88]

- Opções consideradas igualmente atrativas obtenham pontuações coincidentes na escala;
- Uma opção considerada mais atrativa do que outra obtenha uma pontuação superior;
- Se a diferença de atratividade entre duas opções (x, y) for superior à diferença de atratividade entre outras duas opções (w, z), as opções obtenham pontuações de tal modo que a diferença entre as pontuações das opções x e y seja superior à diferença entre as pontuações das opções w e x.

Tendo em conta uma matriz devidamente preenchida com julgamentos consistentes de diferença de atratividade, o MACBETH propõe uma escala numérica de pontuações compatível com os julgamentos efetuados, através da resolução do problema de programação linear (apresentado em detalhe na Secção 4.2.2.). [88]

A solução ótima deste problema é designada por escala MACBETH de base, existindo casos em que a solução ótima não é única e existem múltiplas soluções ótimas, sendo que, nestes casos a escala MACBETH de base proposta é o resultado da média das múltiplas soluções do problema. [91]

A escala MACBETH de base deve posteriormente ser validada, através da comparação gráfica dos tamanhos dos intervalos entre pontuações, podendo estes sofrer ajustes dentro de certos limites sem violar os julgamentos anteriormente introduzidos na matriz, obtendo-se, assim, uma escala de pontuações numérica de intervalos. [88]

4.2.2. Programação Linear MACBETH adaptada ao Estudo

Neste estudo pretendeu-se adaptar a abordagem MACBETH à avaliação de estados de saúde, considerando-se uma abordagem monocritério em que o conjunto de opções em análise consiste num conjunto de diferentes estados de saúde adequadamente descritos.

Neste contexto, seja X um conjunto de estados de saúde finito com $n > 2$ estados, em que $v(x)$ corresponde ao valor atribuído ao estado de saúde x, sendo x^+ e x^- dois elementos do conjunto em que x^+corresponde a um estado tão ou mais atrativo que qualquer outro estado do conjunto e x^- corresponde a um estado que é tão ou menos atrativo do que qualquer outro. Designando-se por C_i, $i = 0, \ldots, 6$, as categorias MACBETH de diferença de atratividade – *nula* (C_0), *muito fraca* (C_1), *fraca* (C_2), *moderada* (C_3), *forte* (C_4), *muito forte* (C_5) e *extrema* (C_6). Considerando o caso mais simples em que a cada par de estados de saúde pode apenas ser associada uma categoria, a escala MACBETH de base, pode ser obtida através da resolução do seguinte problema de programação linear: [88]

$$Min\ [v(x^+) - v(x^-)]$$

Sujeito a:

1. $v(x^-) = 0$ (pontuação arbitrária).
2. $v(x) - v(y) = 0, \forall\, x, y \in C_0$.

3. $v(x) - v(y) \geq i, \forall\, x, y \in C_i$ com i ∈ {1, 2, 3, 4, 5, 6}.
4. $[v(x) - v(y)] - [v(w) - v(z)] \geq i - i', \forall\, x, y \in C_i$ e $\forall\, w, z \in C_{i'}$ com i, i' ∈ {1, 2, 3, 4, 5, 6} e i > i'.

Embora o MACBETH permita que em casos de hesitação ou divergência sejam introduzidas duas ou mais categorias consecutivas, este caso não foi considerado nesta descrição, uma vez que, na adaptação da abordagem a este estudo de avaliação de estados de saúde essa opção não foi apresentada.

4.2.3. Abordagem MACBETH aplicada à Avaliação de Preferências em Saúde

O trabalho desenvolvido nesta tese tem como principal objetivo o desenvolvimento e implementação de uma nova metodologia de avaliação de preferências em saúde recorrendo à abordagem MACBETH, anteriormente descrita. Como tal, começou-se por desenhar um novo protocolo de avaliação de estados de saúde, apresentado esquematicamente na Figura 6.

A nova metodologia proposta divide-se, resumidamente, em quatro etapas principais, sendo que, as duas primeiras etapas recaem, sobretudo, na recolha de informação ordinal e de julgamentos qualitativos de diferenças de atratividade entre elementos, enquanto as duas etapas seguintes consistem na introdução da informação recolhida no *software* M-MACBETH permitindo a obtenção de uma escala de pontuações que reflete a atratividade relativa dos diferentes estados de saúde avaliados.

Na primeira fase é, então, apresentado o conjunto de estados de saúde que se pretende avaliar, recorrendo a um sistema de descrição adequado, e é pedida uma tarefa de ordenação dos mesmos de acordo com a sua atratividade.

Tendo em conta o contexto da avaliação de preferências em saúde e a necessidade de ancorar as valorações obtidas para os diferentes estados de saúde numa escala entre saúde perfeita e morte, no conjunto de estados inicialmente apresentados devem incluir-se os estados de "saúde perfeita" e "morte imediata" que são, posteriormente, definidos como referências superior e inferior, respetivamente. É importante referir, contudo, a possibilidade de existência de estados de saúde considerados como piores do que a morte.

Depois de recolhida a informação ordinal, a fase seguinte consiste num questionamento para obtenção dos julgamentos de diferença de atratividade entre estados de saúde. São, então, apresentados estados de saúde dois a dois e é feita a pergunta "Qual a diferença de atratividade entre os dois estados de saúde apresentados?", cuja resposta deve ser indicada tendo por base a escala semântica MACBETH: n*ula*, *muito fraca*, f*raca*, m*oderada*, *forte*, *muito forte e extrema*. Nesta fase pretende-se recolher julgamentos relativos à diferença de atratividade entre pares de estados de saúde ordenados consecutivamente e, ainda, dois julgamentos adicionais, relativos à diferença de atratividade entre o estado mais atrativo e uma opção intermédia e entre a mesma opção intermédia e o estado menos atrativo, de modo a testar a consistência dos julgamentos.

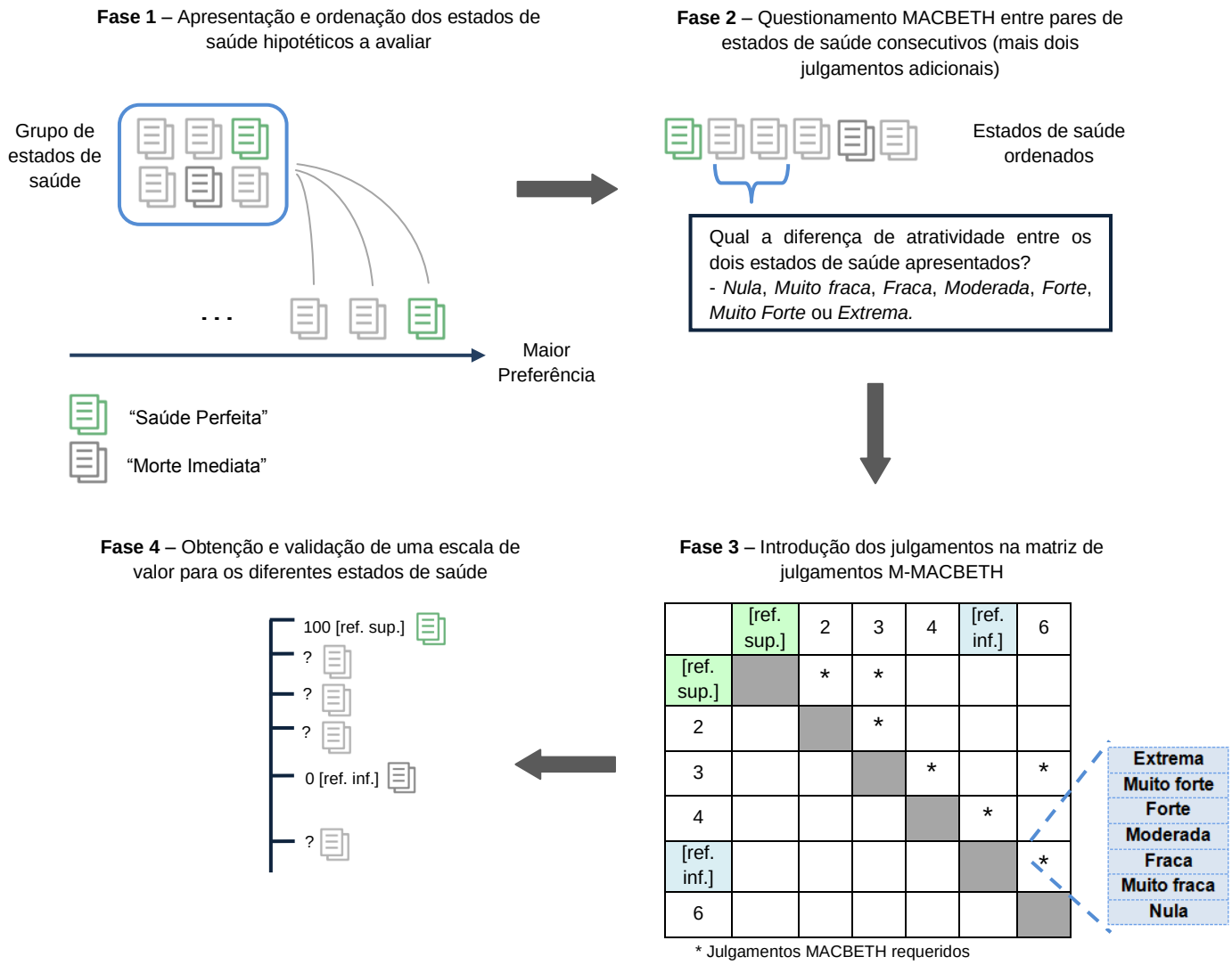

Figura 6. Representação esquemática dos diferentes passos necessários para avaliação de estados de saúde recorrendo ao método proposto, baseado na abordagem MACBETH.

À medida que se vão obtendo os julgamentos estes são introduzidos no *software* M-MACBETH, preenchendo-se a matriz de julgamentos de acordo com a representação esquemática da Figura 6 - Fase 3, onde as opções se encontram previamente ordenadas de acordo com as respostas obtidas na tarefa de ordenação. Nesta fase são, também, definidas a referência superior "saúde perfeita" e a referência inferior "morte imediata", às quais são atribuídos os valores de 100 e 0, respetivamente. No caso de não serem definidas explicitamente as referências o M-MACBETH atribui, por defeito, o valor 100 à opção mais atrativa e o valor 0 à opção menos atrativa. [89]

Por fim, é possível obter uma escala termométrica de pontuações que reflete a atratividade relativa dos estados de saúde em análise, sendo que esta pode ser ajustada dentro de certos limites, respeitando os julgamentos introduzidos.

4.3. Desenvolvimento de Questionário de Avaliação de Estados de Saúde

Uma vez desenvolvido o desenho de um novo protocolo de avaliação de preferências em saúde baseado na abordagem MACBETH pretendeu-se implementar e testar o mesmo, bem como, analisar congruências entre o novo método proposto e uma técnica de valoração tradicional, baseada em julgamentos numéricos. Com o objetivo de implementar ambos os métodos referidos, bem como, recolher dados relativos à preferência, ao grau de numeracia e, ainda, de caraterização sociodemográfica e do estado de saúde atual da população, propõe-se a implementação de um questionário *online* recorrendo-se para tal à plataforma Qualtrics.

A opção pelo desenvolvimento de um questionário *online* para recolha de dados justificou-se pelo fato de esta abordagem permitir uma distribuição fácil e rápida do mesmo, em que os indivíduos inquiridos podem responder de acordo com a sua conveniência de tempo uma vez que não é necessária a presença de um entrevistador, evitando-se também, deste modo, a influência deste nas respostas obtidas. O recurso à implementação de questionários via Web em estudos de diversas áreas, em particular em estudos na área da elicitação de preferências em saúde é uma abordagem inovadora em desenvolvimento. [92] [93] [94] Este desenvolvimento é motivado, sobretudo, pelos inúmeros desafios logísticos e limitações de recursos associados aos métodos convencionais de elicitação de preferências, nomeadamente à condução de entrevistas pessoais individuais e ainda pelo fato de que, cada vez mais, as novas tecnologias e a internet são uma realidade acessível à maior parte da população nos países desenvolvidos. As vantagens enunciadas e o fato de o recurso a plataformas Web para implementação de questionários ser uma abordagem inovadora já com precedentes na área contribuíram para esta escolha.

As várias etapas subjacentes ao desenvolvimento e implementação deste questionário *online* são detalhadamente descritas nas secções seguintes.

4.3.1. Selecção do Método Numérico

O método convencional de valoração de estados de saúde, baseado em julgamentos numéricos, selecionado para apresentar conjuntamente com o novo método proposto foi o *Time Trade-Off*, anteriormente descrito na Secção 3.1.3.

A opção pelo método *Time Trade-Off* recaiu sobre o fato de este ser um método de valoração de estados de saúde dominante no campo da economia de saúde, sendo, hoje em dia, amplamente aplicado para valoração de estados de saúde em diferentes países e populações. [95] Adicionalmente, o método *Time Trade-Off* foi, também, utilizado para estimar o conjunto de valores do instrumento genérico EQ-5D no caso português, sendo que este instrumento genérico de medição de QVRS é o mais comumente utilizado hoje em dia, nomeadamente nas avaliações efetuadas pelo NICE. [76]

4.3.2. Seleção dos Estados de Saúde

A descrição dos diferentes estados de saúde que se pretendem avaliar tem como base o sistema classificativo do EQ-5D. Esta decisão funde-se com a decisão relativa à escolha do método convencional a implementar apresentada na secção anterior.

Segundo esta classificação cada estado de saúde é descrito de acordo com cinco dimensões (mobilidade, cuidados pessoais, atividades habituais, dor/mal-estar, ansiedade/depressão) e cada dimensão apresenta três níveis possíveis (1 – sem problemas, 2 – alguns problemas, 3 – problemas severos). Cada estado de saúde é, assim, descrito pela atribuição de um nível a cada dimensão, por exemplo, o estado 11111 representa o melhor estado de saúde, sem problemas em nenhuma das dimensões, e o estado 33333 representa o pior estado de saúde, com problemas severos em todas as dimensões. [62]

Tendo em conta o sistema de caraterização de estados de saúde apresentado podem definir-se 243 (3^5) estados de saúde distintos, contudo, nem todos os estados vão ser incluídos neste estudo. Foram, então, selecionados para avaliação apenas os mesmos 24 estados de saúde considerados no estudo de valoração do EQ-5D em Portugal e ainda os estados 11111, 33333 e "morte imediata", de modo a ser possível a posterior comparação dos valores obtidos neste estudo com os valores anteriormente obtidos na valoração dos estados do EQ-5D para a população portuguesa. [76]

Tal como no estudo de valoração do EQ-5D em Portugal e tendo em conta, que um indivíduo não consegue avaliar num único questionário todos os estados de saúde seleccionados, estes foram divididos em quatro grupos de acordo com a Tabela 4. [76] [96] Assim, a cada indivíduo é atribuído aleatoriamente um dos quatro grupos, constituído por seis estados de saúde diferentes e ainda pelo estado 33333, comum a todos os grupos. A distribuição dos estados pelos diferentes grupos teve em conta o grau de severidade destes, de modo a assegurar que em cada grupo se encontram estados classificados como muito ligeiros, ligeiros, moderados e severos.

Tabela 4. Distribuição dos estados de saúde por grupo de acordo com o seu grau de severidade.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
13311 ***	12111 *	11113 **	21111 *
22222 ***	11131 **	32313 ***	23232 ****
11112 *	32211 ***	11211 *	11121 *
11133 **	21323 ***	22121 **	11312 **
32223 ****	22233 ****	13332 ****	33323 ****
33321 ****	23313 ****	33232 ****	22122 ***
33333	33333	33333	33333

* estado muito ligeiro, ** estado ligeiro, *** estado moderado, **** estado severo

4.3.3. Estruturação do Questionário

Atendendo ao fato de que se pretende implementar o questionário *online* e à constatação da existência de dificuldades em obter respostas a este tipo de questionários quando é requerida aos inquiridos uma elevada carga temporal, procurou-se, dentro do possível, sintetizar o questionário de modo a recolher apenas as informações fundamentais para atingir os objetivos deste estudo.

Assim, de acordo com os principais objetivos definidos neste trabalho e com as hipóteses que se pretendem estudar, o questionário de avaliação de estados de saúde foi estruturado em cinco módulos, conforme apresentado na Figura 7, cada um com o intuito de recolher informações de natureza diferente.

Figura 7. Estruturação do questionário em cinco módulos, cada um correspondente à recolha de diferentes tipos de dados.

O primeiro módulo pretende caraterizar a saúde atual dos indivíduos inquiridos através do preenchimento do questionário EQ-5D, um instrumento genérico de avaliação da qualidade de vida relacionada com a saúde. Considerou-se de interesse a apresentação desta tarefa em primeiro lugar, uma vez que, permite um primeiro contacto com o sistema descritivo do EQ-5D utilizado na caraterização dos estados de saúde hipotéticos que se pretendem avaliar.

De seguida são recolhidos os julgamentos numéricos associados ao método *Time Trade-Off* e os julgamentos semânticos associados à abordagem MACBETH, sendo a ordem de apresentação destas tarefas, em cada questionário, aleatória. Uma descrição pormenorizada das várias etapas associadas a cada um destes métodos é apresentada nas Secções 4.3.4.2 e 4.3.4.3.

Com o intuito de dar resposta às hipóteses anteriormente levantadas relativamente à numeracia, foi incluído um módulo que consiste na aplicação de um teste de avaliação das competências numéricas dos indivíduos. Embora, também fosse de interesse incluir neste módulo um teste às competências verbais dos indivíduos, tendo em consideração a dimensão do questionário optou-se por apenas selecionar um teste de avaliação da numeracia. Esta decisão é justificada pelo fato de a evidência encontrada apontar para uma influência da numeracia na valoração de estados de saúde recorrendo a métodos convencionais, influência esta que se considerou de interesse principal averiguar.

O módulo seguinte visa a identificação do modo preferencial de cada indivíduo expressar julgamentos de valor, em números ou palavras, pretendendo-se, ainda, no âmbito desta secção incluir um espaço de resposta aberta que permite aos indivíduos inquiridos comentar a preferência manifestada.

Por último foi introduzido um módulo de caraterização sócio-demográfica, considerado aspetos como a idade, o género, o nível de escolaridade, o estado civil, entre outros.

4.3.4. Implementação do Questionário *online* na Plataforma Qualtrics

Uma vez definida a estrutura geral do questionário a desenvolver, é de seguida descrita de forma detalhada a implementação efetuada para cada uma dos módulos definidos anteriormente, assim como, algumas adaptações necessárias tendo em conta as ferramentas disponíveis na plataforma Qualtrics. É relevante referir que esta plataforma tem disponíveis ferramentas de aleatorização, tendo-se feito uso destas relativamente à ordem de apresentação das tarefas de atribuição de julgamentos numéricos e não numéricos e relativamente aos estados de saúde apresentados para avaliação. Esta plataforma permite, ainda, modificações de *layout* e de estilo do questionário desenvolvido através da utilização de código CSS, pelo que se recorreu a esta ferramenta para ajustar as questões pré-formatadas disponíveis.

4.3.4.1. Caraterização do Estado de Saúde Atual

Com o objetivo de proceder à caraterização do estado de saúde atual dos indivíduos inquiridos foi implementado o instrumento genérico EQ-5D, anteriormente descrito na Secção 3.2.3., tendo-se como base a versão portuguesa do mesmo (Figura 21, Anexo A). [97]

Este instrumento consiste essencialmente em duas fases: uma escala VAS e um sistema descritivo que avalia as cinco dimensões de saúde (mobilidade, cuidados pessoais, atividades habituais, dor/mal-estar, ansiedade/depressão) numa escala com três níveis – sem problemas, alguns problemas e problemas severos. A escala VAS pretende que o indivíduo indique o valor que atribui ao seu estado de saúde atual numa escala entre 0 e 100, correspondentes respetivamente ao pior e melhor estados de saúde imagináveis, enquanto o sistema descritivo permite obter uma combinação de índices que definem o estado de saúde de entre 243 estados possíveis.

Adicionalmente o EQ-5D inclui, ainda, uma questão de comparação do nível de saúde do inquirido durante os últimos doze meses com o seu estado de saúde atual, onde os níveis de resposta possíveis são: “melhor”, “o mesmo”, “pior”.

No que diz respeito à implementação do EQ-5D na plataforma Qualtrics, importa referir que a escala VAS foi apresentada na horizontal, tendo em conta o formato de questão disponível, no entanto, esta alteração foi analisada noutros estudos da área onde se concluiu que não conduz a nenhum tipo de viés significativo. [98]

4.3.4.2. Recolha de Julgamentos de Valor Numéricos – *Time Trade-Off*

O método de *Time Trade-Off*, de acordo com a implementação seguida no estudo de valoração dos estados de saúde do EQ-5D em Portugal [76], pode dividir-se em duas etapas distintas de acordo com o tipo de exercício apresentado aos indivíduos inquiridos.

A primeira etapa associada ao método TTO consiste na indicação se o estado de saúde a ser avaliado é considerado pelo respondente melhor ou pior do que a morte. No caso de o estado de saúde ser indicado como sendo melhor do que a morte é iniciado um processo iterativo, em que o

indivíduo pode escolher entre duas opções: viver no estado apresentado para avaliação por um período de 10 anos (seguidos de morte imediata) ou viver no estado de saúde perfeita (11111) durante t anos, sendo que este valor vai sendo alterado (começando-se em 5 anos e aumentando até 10 anos ou diminuindo até 0 anos), até ser identificado o ponto de indiferença entre as duas opções apresentadas. O valor a atribuir ao estado em avaliação é obtido a partir do ponto de indiferença determinado através da expressão $h = t/10$. [76]

No caso de o estado de saúde a ser avaliado ser considerado como sendo pior do que a morte, é igualmente pedido ao indivíduo que escolha entre duas opções: morrer imediatamente ou viver algum tempo, t anos, no estado de saúde apresentado para avaliação seguidos por $(10 - t)$ anos no estado de saúde perfeita (11111). Mais uma vez, o valor a atribuir ao estado de saúde em avaliação é obtido a partir do ponto de indiferença entre as duas opções apresentadas, através da expressão $h = (-10 + t)/t$. [76]

Habitualmente, este método é implementado com recurso a entrevistas pessoais e com o auxílio de apoios visuais que ajudam os indivíduos na realização e compreensão das tarefas requeridas. No entanto, este modelo de implementação apresenta uma carga temporal elevada sendo estimada que a valoração de seis estados demore aproximadamente 20 minutos [99], pelo que, embora inicialmente se tenha realizado uma tentativa de implementação *online* de um protocolo semelhante ao habitualmente utilizado nas entrevistas pessoas, isto é, recorrendo a esquemas visuais e ao processo iterativo descrito anteriormente, optou-se por não avançar com esse modelo.

Numa tentativa de simplificar o questionário optou-se, assim, por seguir uma adaptação simplificada da metodologia anteriormente descrita. Como tal, foi mantida a primeira etapa de determinação se o estado de saúde a ser avaliado é considerado pelo respondente melhor ou pior do que a morte, sendo apresentada para cada estado de saúde uma questão semelhante ao exemplo da Figura 8. Nesta questão importa evidenciar, que foi dada indicação aos indivíduos que cada estado de saúde deve ser considerado com uma duração de 10 anos, sem sofrer qualquer alteração, seguidos por morte imediata. [96]

Relativamente à recolha dos julgamentos numéricos de acordo com o método TTO, foram desenvolvidos dois tipos de questão consoante o estado de saúde foi considerado como sendo melhor ou pior do que a morte. Assim, quando na primeira etapa relativa ao método TTO, em relação a um determinado estado de saúde, é selecionada a resposta “Viver 10 anos no estado apresentado” a questão seguinte apresentada para esse mesmo estado é a questão definida para estados considerados melhores do que a morte. Pelo contrário, quando a resposta selecionada é “Morte imediata”, a questão seguinte apresentada relativamente a esse mesmo estado é a questão desenvolvida para estados considerados piores do que a morte. Quando um estado é considerado igual à morte, nenhuma questão é apresentada, uma vez que o valor atribuído a esse estado é 0.

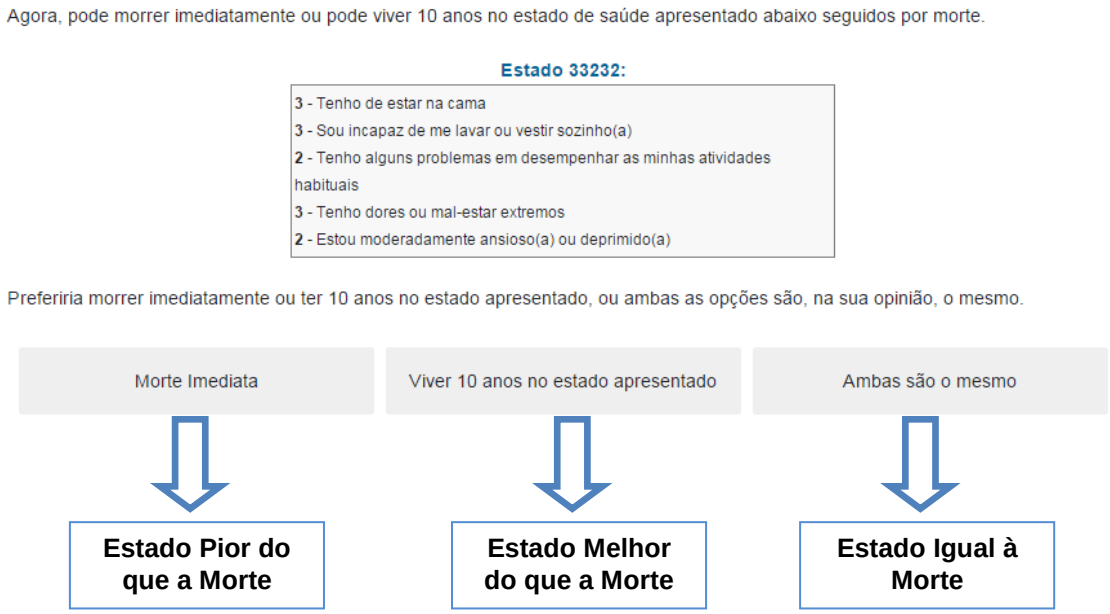

Figura 8. Exemplo de questão de avaliação se um dado estado de saúde é considerado melhor ou pior do que morte imediata.

Para cada estado de saúde considerado melhor do que a morte foi, assim definida uma questão semelhante à apresentada na Figura 9, tendo como base o método TTO seguido noutros estudos. [100] [101] [102] São, então, apresentadas ao indivíduo duas opções, tal como no método referido anteriormente, no entanto, não é seguido um processo iterativo até se atingir o ponto de indiferença entre estas, sendo este processo simplificado através da apresentação direta de uma escala horizontal, limitada pelos valores 0 e 10, que representa o número de anos no estado de saúde perfeita e a qual se pede que o indivíduo ajuste, indicando diretamente o ponto de indiferença entre as opções.

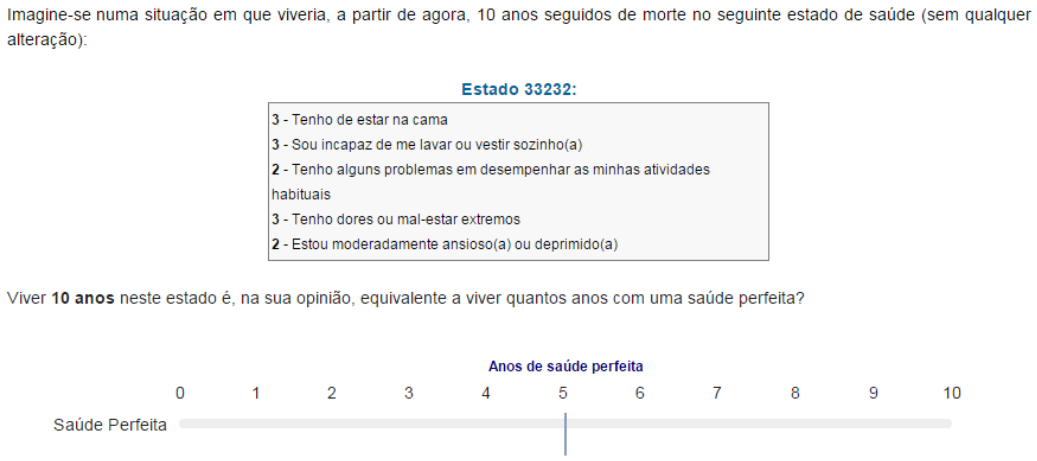

Figura 9. Exemplo de questão de recolha de julgamentos para estados considerados melhores do que a morte.

Para estados considerados piores do que a morte foi seguida uma estrutura semelhante à apresentada para estados considerados melhores do que a morte, alternando-se, contudo, as duas opções apresentadas. Na Figura 10 é apresentado um exemplo de questão para este caso.

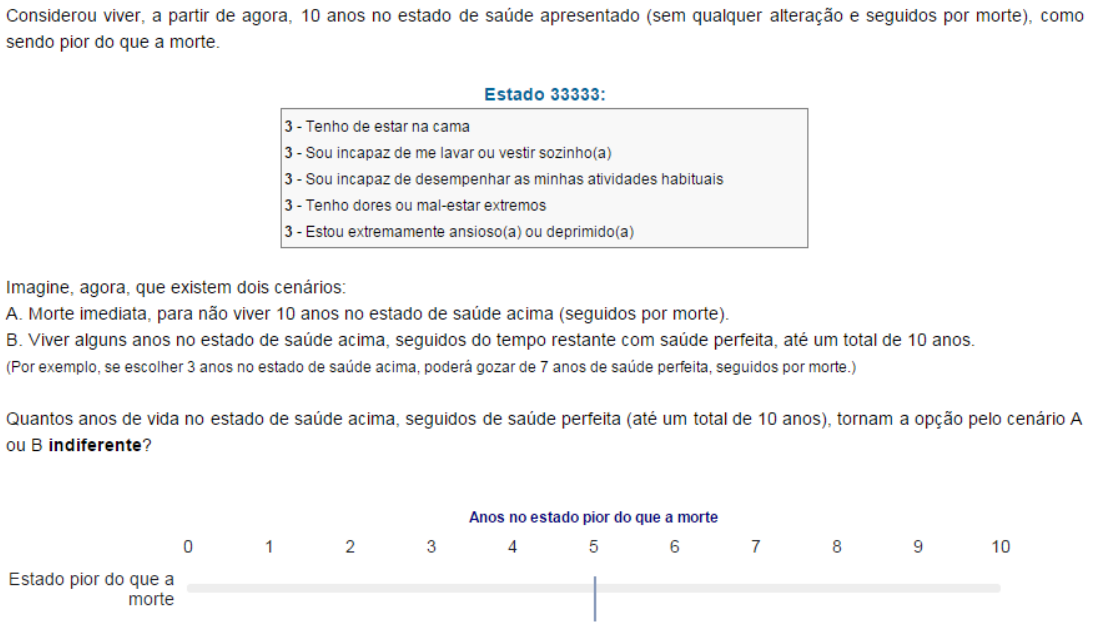

Figura 10. Exemplo de questão de recolha de julgamentos para estados considerados piores do que a morte.

4.3.4.3. Recolha de Julgamentos de Valor Não Numéricos – MACBETH

Tendo em conta a estruturação do novo protocolo de avaliação de estados de saúde recorrendo à abordagem MACBETH, apresentado anteriormente na Secção 4.2.3., a implementação desta abordagem recorrendo à Plataforma Qualtrics foi dividida em duas etapas: uma primeira etapa de apresentação e ordenação do conjunto de estados de saúde apresentados e uma segunda etapa correspondente à recolha de julgamentos semânticos de diferença de atratividade entre os pares de estados de saúde previamente definidos.

Devido a limitações da plataforma escolhida para implementação *online* do questionário no que diz respeito à tarefa de ordenação foi necessário seguir um tipo de implementação que permitisse o posterior reconhecimento da lista ordenada de estados de saúde, essencial para a etapa seguinte desta metodologia. Como tal, de modo a obter uma lista ordenada dos vários estados foi pedida uma seleção sucessiva do melhor estado de entre o conjunto de estados apresentados, sendo que, após ser selecionado um estado este é eliminado do conjunto e assume uma posição no *ranking* de acordo com o número da pergunta em que foi selecionado, o processo continua até que todos os estados tenham uma posição definida no *ranking*.

Importa referir, que se optou por não apresentar na lista para ordenação o estado de saúde perfeita, uma vez que, é considerado que para qualquer individuo que tenha compreendido totalmente a tarefa pedida esse é, sem dúvida, o primeiro estado a selecionar. Como tal, o melhor

estado de saúde selecionado na primeira pergunta corresponde ao estado classificado em segundo lugar no *ranking* de ordenação por ordem de preferência.

A segunda etapa de implementação da nova metodologia proposta consistiu na apresentação de pares de estados de saúde e na recolha de julgamentos de diferença de atratividade entre estes, de acordo com a Figura 11. Os julgamentos recolhidos foram os determinados na Secção 4.2.3., isto é, julgamentos de diferenças de atratividade entre pares de estados de saúde ordenados consecutivamente e adicionalmente um julgamento entre o estado de saúde perfeita e um estado intermédio, o estado em terceiro lugar na lista ordenada, e outro julgamento entre este e o estado que ocupa a ultima posição da lista ordenada.

Importa referir que não foi permitida a seleção de julgamentos múltiplos em caso de divergência ou hesitação.

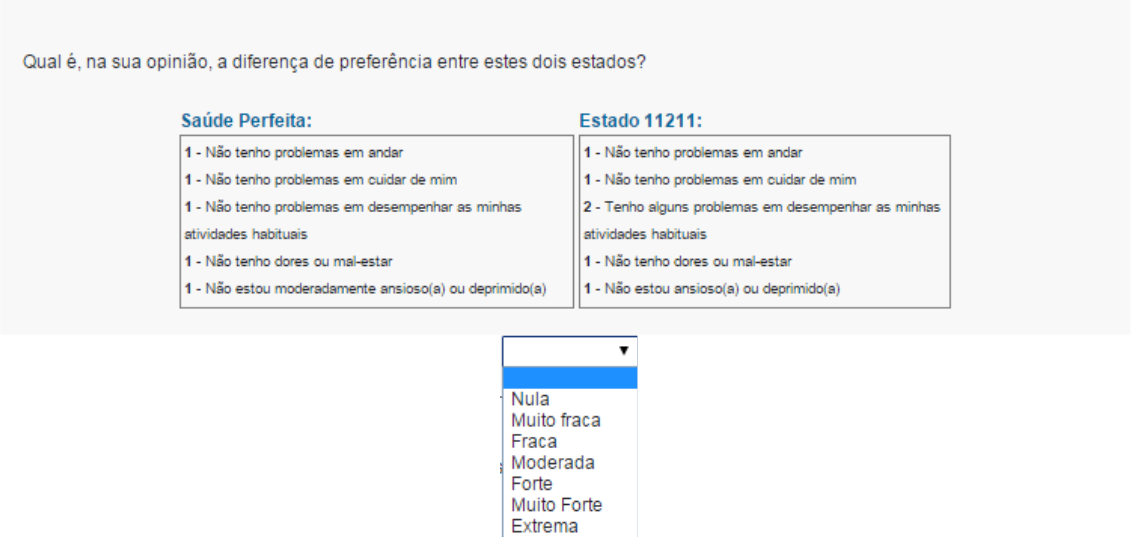

Figura 11. Exemplo de questão de recolha de julgamentos de diferença de atratividade com base na escala semântica MACBETH.

4.3.4.4. Avaliação das Preferências

No que diz respeito à avaliação da preferência dos indivíduos inquiridos relativamente ao método de expressar julgamentos de valor preferencial, números ou palavras, foram implementados dois tipos de questões. Uma questão em que se pede aos indivíduos que escolham o método preferencial de entre “Julgamentos Semânticos”, “Julgamentos Quantitativos” ou “Indiferente”. Sendo que, tanto a opção julgamentos semânticos como a opção julgamentos quantitativos são acompanhadas por uma representação gráfica ilustrativa do tipo de questão que lhe está associada. A segunda questão consistiu num espaço de resposta aberta onde se pediu aos inquiridos que justificassem a escolha anteriormente efetuada.

4.3.4.5. Avaliação da Numeracia

A numeracia foi avaliada através de três questões validadas, anteriormente utilizadas noutros estudos de determinação da influência da numeracia na obtenção de utilidades de saúde, sendo a cada indivíduo atribuída uma pontuação de acordo com o número de respostas corretas, tendo-se considerado não respostas como sendo respostas erradas. [10] [11]

As questões apresentadas foram:

1 – "Em 1000 lançamentos de uma moeda ao ar, quantas vezes é esperado que saia coroa?"

2 – "Quanto é um por cento de 1000?"

3 – "Um é que percentagem de 1000?"

A primeira questão encontra-se, assim, relacionada com probabilidades e pretende avaliar o entendimento básico destas de cada indivíduo. Na segunda questão é pedido aos indivíduos que convertam uma percentagem numa proporção e na terceira questão é pedido que convertam uma proporção numa percentagem.

Embora outros testes de avaliação de numeracia tenham sido considerados, nomeadamente, o *Subjective Numeracy Scale* (SNS) [103] que consiste numa medida de autoavaliação da capacidade para desenvolver tarefas matemáticas com uma componente relativa à preferência do uso de informações numéricas ou não numéricas, e ainda o teste de Lipkus *et al.* (2001) [104], entre outros, optou-se pelas três questões anteriormente referidas por uma questão de simplicidade e de não sobrecarregar o questionário.

4.3.5. Estratégia de Distribuição e População-Alvo

A implementação e a aplicação de questionários implicam uma definição prévia de dois aspetos fundamentais que se correlacionam entre si: identificação da população-alvo do mesmo e definição da estratégia de distribuição a seguir.

A população-alvo definida para este estudo consistiu numa amostra de indivíduos da população geral portuguesa com idade superior a 18 anos. Os indivíduos foram recrutados para participar neste estudo, respondendo ao questionário, através de um convite enviado por *email* no qual foi sucintamente explicado o propósito do mesmo e indicado um *link* que remetia para o questionário e para o seu preenchimento. No *email* enviado foi, ainda, solicitado o reenvio do mesmo a outros contactos, seguindo-se, deste modo, uma estratégia de amostragem não probabilística – amostragem por redes ou bola de neve. [105] A etapa inicial da amostragem bola de neve aplicada envolveu a distribuição de convites à participação no estudo a uma amostra de indivíduos, da rede de contatos pessoais dos elementos envolvidos no mesmo, constituída por 54 indivíduos.

4.4. Obtenção de Escalas de Valor

Uma vez implementado o questionário proposto e concluída a fase de recolha de dados de acordo com a estratégia anteriormente descrita, a etapa seguinte consistiu na valoração dos estados

de saúde propriamente dita, tendo em conta os julgamentos numéricos e não numéricos recolhidos de acordo com cada um dos métodos. Os processos de valoração dos diferentes estados de saúde são descritos nas secções seguintes.

4.4.1. Valoração *Time Trade-Off*

No que diz respeito aos julgamentos de valor numéricos recolhidos de acordo com o método TTO, os valores para os diferentes estados de saúde são obtidos de acordo com o processo descrito na Secção 4.3.4.2., ou seja, para estados de saúde considerados como melhores do que a morte o valor que lhes é atribuído é dado por $h = t/10$, enquanto para estados considerados como sendo piores do que a morte o valor atribuído é dado pela expressão $h = (-10 + t)/t$, sendo que para ambos os casos, o valor de t diz respeito ao número de anos que torna a escolha entre as duas alternativas apresentadas indiferente.

Tendo em conta que a escala apresentada para indicação do valor de t permite uma variação entre 0 e 10 anos com uma precisão de 0.1, os valores derivados para os diferentes estados de saúde podem variar entre -99 e 1. Assim, de modo a tornar os valores obtidos consistentes com aqueles obtidos em estudos anteriores da área, nomeadamente com o estudo de valoração dos estados do EQ-5D em Portugal [76], os valores negativos (valores obtidos para estados considerados como piores do que a morte) foram ajustados para uma escala entre -1 e 0, recorrendo-se, para tal, à transformação monotónica $h' = h/(1 - h)$, amplamente utilizada noutros estudos. [96] [76] A escala de valores positivos (valores obtidos para estados considerados como melhores do que a morte) não sofreu ajustes, uma vez que varia entre 0 e 1.

Os valores de TTO obtidos variam, assim, entre -1 e 1, onde 1 representa a saúde perfeita e 0 representa a morte imediata, sendo os valores negativos atribuídos a estados considerados como sendo piores do que a morte.

4.4.2. Valoração MACBETH

O processo de valoração dos diferentes estados de saúde de acordo com o novo método proposto tendo como base a abordagem MACBETH foi apresentado anteriormente na Secção 4.2. e é esquematicamente representado, aplicado ao caso da avaliação de estados de saúde, na Figura 12. A primeira fase deste processo consiste na criação de um nó critério associado a cada indivíduo inquirido, seguindo-se a identificação das referências superior e inferior ("saúde perfeita" e "morte imediata", respetivamente), bem com, das diferentes opções em análise, neste caso os diferentes estados de saúde em avaliação. De seguida, os diferentes estados de saúde apresentados na matriz de julgamentos são ordenados de acordo com as respostas obtidas na tarefa de ordenação anteriormente descrita e, por fim, são introduzidos na matriz ordenada os julgamentos semânticos recolhidos, permitindo-se, assim, a obtenção de uma escala numérica MACBETH no caso de os julgamentos introduzidos serem consistentes.

Tendo em conta que, de acordo com o modelo de recolha de dados definido neste estudo, a recolha de julgamentos e a introdução dos mesmos no *software* M-MACBETH são duas tarefas realizadas em momentos distintos, não foi possível a apresentação a cada indivíduo da escala MACBETH obtida, tendo em conta os julgamentos efetuados, para validação da mesma ou ajuste dos intervalos. Considerou-se, assim, automaticamente válida a escala proposta pelo *software*.

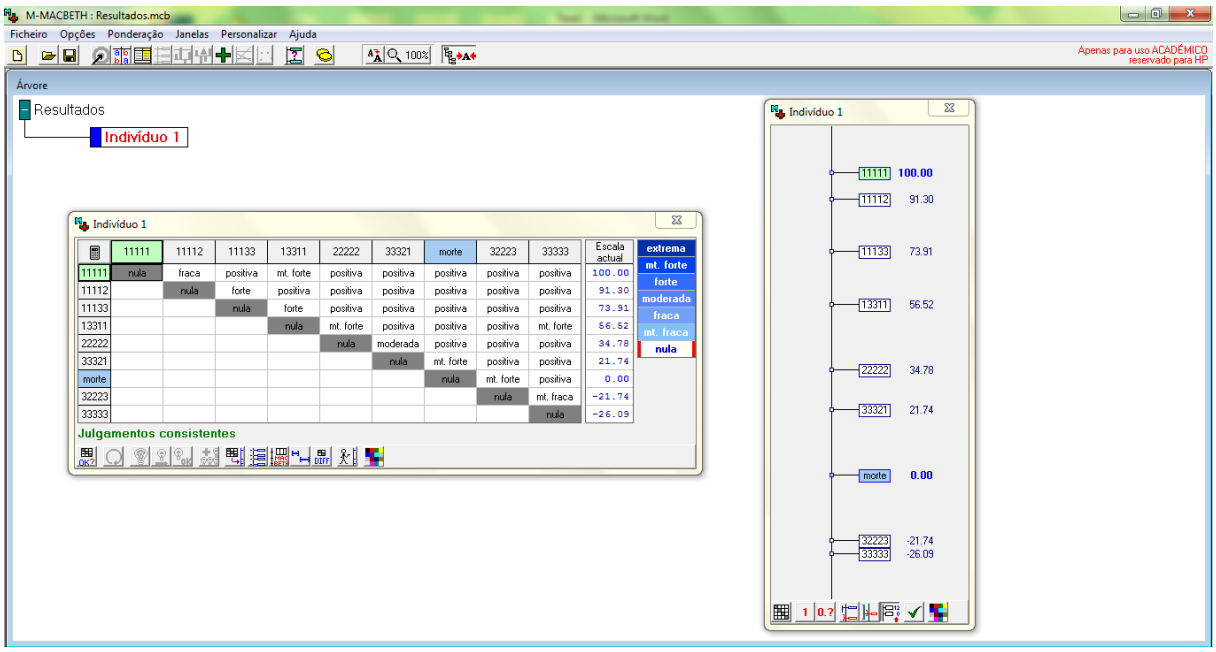

Figura 12. Aplicação do MACBETH para valoração do conjunto de estados de saúde apresentado ao Indivíduo 1.

Tal como, descrito anteriormente para as escalas obtidas recorrendo ao método TTO, também as escalas obtidas recorrendo ao MACBETH foram ajustadas, de modo à obtenção de escalas de valores consistentes com outros estudos da área, ou seja, escalas de valores entre -1 e 1, em que 1 representa a saúde perfeita e 0 representa a morte imediata, sendo os valores negativos atribuídos a estados considerados como sendo piores do que a morte.

Numa primeira fase foi, então, aplicada a todos os valores a transformação linear $h' = h/100$, de modo a obter uma escala limitada superiormente pelo valor 1. Numa segunda fase, tendo em conta que tal como para o TTO, a escala de valores negativos não é limitada inferiormente pelo valor -1, foi necessário a aplicação da transformação monotónica $h' = h/(1-h)$.

4.5. Critérios de Exclusão

Após a recolha de dados e simultaneamente à valoração dos diferentes estados de saúde de acordo com ambos os métodos, verificou-se que alguns indivíduos não responderam à totalidade do questionário, tendo desistido a meio e, como tal, estes foram eliminados da análise. Adicionalmente algumas das respostas obtidas revelaram julgamentos inconsistentes ou incompreensão da tarefa pedida. Definiram-se, assim, os seguintes critérios de exclusão objectivos relativamente aos julgamentos obtidos associados ao método MACBETH:

1) Inconsistência lógica de ordenação dos estados, isto é, um estado de um par que é melhor do que o outro em pelo menos uma dimensão e não é pior em nenhuma é selecionado depois do segundo na tarefa de ordenação;
2) Todos os estados foram considerados como piores do que a morte;
3) Obtenção de matriz de julgamentos MACBETH inconsistente.

Relativamente aos julgamentos recolhidos de acordo com o método TTO, foram também definidos critérios de exclusão associados à não compreensão da tarefa pedida, estando estes de acordo com critérios anteriormente usados noutros estudos de aplicação do método para valoração de estados de saúde:

1) Todos os estados foram considerados como piores do que a morte [76];
2) O mesmo valor de TTO foi atribuído a todos os estados [76];
3) Para estados considerados melhores do que a morte foi selecionado o valor $t = 0$ [100];
4) Para estados considerados piores do que a morte foi selecionado o valor $t = 0$.

Ainda relativamente ao método TTO foram analisados dois critérios adicionais baseados na obtenção de inconsistências lógicas de ordenação e na obtenção de inconsistências lógicas de ordenação graves (quando a diferença entre os valores atribuídos a dois estados ordenados inconsistentemente é superior a 0.5). [76] Contudo, optou-se por não considerar estes dois critérios como critérios objetivos de exclusão, contrariamente ao que foi feito relativamente ao método MACBETH, uma vez que, neste método não é apresentada diretamente uma tarefa de ordenação, resultando a ordenação final dos estados das respostas que vão sendo obtidas e, como tal, garantir a consistência da ordenação implicaria um individuo ter presente em todos os momentos as respostas anteriormente dadas.

No que diz respeito às exclusões efetuadas de acordo com os critérios definidos, sempre que se verificou num questionário pelo menos um dos problemas referidos em ambos os métodos, este questionário foi excluído da análise. Por outro lado, quando apenas um dos métodos apresenta um ou mais dos problemas referidos apenas este é excluído da análise.

4.6. Análise de Dados Recolhidos

Uma vez concluída a fase de recolha dos dados e a valoração dos estados de saúde tendo em conta os julgamentos numéricos e não numéricos recolhidos em cada questionário prosseguiu-se para a etapa final deste estudo, a análise dos dados recolhidos tendo como objetivo dar resposta às questões anteriormente definidas: descrição do processo de obtenção da amostra populacional válida para análise; caraterização da amostra populacional relativamente às variáveis sociodemográficas, ao estado de saúde atual e a grau de numeracia; comparação entre os dois métodos implementados; caraterização das preferências e, por último, estudo da influência da numeracia nas preferências manifestadas, nas escalas de valores obtidas e no número de exclusões efetuadas.

A análise dos dados recolhidos teve como base, maioritariamente a utilização de ferramentas de análise estatística descritiva e comparativa, sendo estas descritas na Secção 4.6.1., contudo, outro tipo de análise foi ainda considerada de interesse a qual é descrita na Secção 4.6.2.

4.6.1. Análise Estatística

A análise estatística dos dados obtidos pode organizar-se em duas fases distintas, em que uma primeira fase consiste na descrição sistemática destes recorrendo-se a estatísticas descritivas univariadas, enquanto a segunda fase consiste na aplicação de análises estatísticas descritivas bivariadas e análises estatísticas inferenciais, nomeadamente testes de hipóteses.

Assim, relativamente à primeira fase referida de análise descritiva utilizaram-se, para caraterizar as variáveis numéricas, medidas estatísticas de tendência central e variabilidade, tais como, a média e o desvio-padrão. No caso de algumas variáveis numéricas revelou-se de interesse a sua conversão em variáveis categóricas, sendo que todas as variáveis deste tipo foram descritas pela apresentação de frequências absolutas e relativas.

Este tipo de análise foi aplicado na descrição do processo de obtenção da amostra populacional válida para análise, caraterização das amostras populacionais obtidas e de vários subgrupos amostrais de interesse no que diz respeito às variáveis sociodemográficas, ao estado de saúde atual e ao grau de numeracia, caraterização das escalas de valores obtidas e caraterização das preferências.

Na segunda fase da análise de dados, para analisar diferenças estatisticamente significativas entre diferentes subgrupos amostrais recorreram-se a testes paramétricos (Teste t e ANOVA) e testes não paramétricos (Teste χ^2, Teste exato de Fisher ou Teste de Mann-Whitney). De modo a identificar e analisar associações ou correlações entre variáveis recorreu-se ao teste do χ^2 da independência no caso das variáveis serem nominais, quando as variáveis são quantitativas e cumprirem os pré-requisitos da estatística paramétrica recorreu-se à correlação de Pearson (r), por outro lado, quando tal não se verifica recorreu-se à correlação de Spearman (ρ). Para todas as análises efetuadas considerou-se um nível de significância de $p \leq 0.05$.

Adicionalmente, para o caso específico do estudo da influência da numeracia nas preferências manifestadas recorreu-se a uma regressão logística multinominal, definindo-se como variável dependente a preferência manifestada pelos indivíduos que pode assumir três categorias diferentes ("Julgamentos semânticos", "Julgamentos numéricos" ou "Indiferente") e como variável independente definiu-se o grau de numeracia dos indivíduos, sendo esta uma variável numérica a variar entre 0 e 3.

Para manipulação da base de dados, bem como para todas as análises estatísticas efetuadas recorreu-se ao *software* R, versão 3.1.3.

4.6.2. Outras Análises

Adicionalmente às análises estatísticas referidas na secção anterior, revelou-se de interesse recorrer a análises adicionais, nomeadamente ao uso de uma das funcionalidades do *software* M-MACBETH que permite o ajuste da escala proposta com base nos julgamentos semânticos introduzidos na matriz de julgamentos a valores específicos, caso estes sejam compatíveis com a matriz de julgamentos introduzida. Assim, para uma amostra dos indivíduos da amostra populacional

válida que atribuíram julgamentos válidos de acordo com ambos os métodos procedeu-se à tentativa de ajuste da escala MACBETH proposta com a escala TTO obtida.

Outra análise realizada consistiu na determinação das diferenças entre ambas as escalas de valores médios obtidas e na identificação de diferenças absolutas superiores a 0.1 e a 0.05, que correspondem na escala TTO a variações de 1 ano e 6 meses, respetivamente.

5. Resultados

Uma vez concluídas as fases de estruturação, desenvolvimento e implementação da metodologia proposta, apresentam-se neste capítulo os resultados obtidos, nomeadamente os resultados da aplicação das várias análises anteriormente definidas aos dados recolhidos através do questionário de avaliação de estados de saúde proposto.

Na primeira secção do presente capítulo começa-se por apresentar uma descrição detalhada do processo de obtenção da amostra válida para análise, descrevendo-se o número de exclusões associadas a cada um dos critérios de exclusão definidos na Secção 4.5. Tendo em conta, a aleatorização da ordem de apresentação das tarefas de atribuição de julgamentos semânticos e numéricos, analisou-se ainda, nesta secção, a possível influência da ordem de apresentação destas tarefas, tanto no número de exclusões obtidas como, também, no número de questionários incompletos obtido.

A secção seguinte consiste na caraterização da amostra populacional deste estudo, antes e após exclusões, no que diz respeito às variáveis sociodemográficas, ao estado de saúde atual e ao grau de numeracia e, ainda, na comparação entre estas de modo a identificar a existência de algum padrão relativamente às exclusões efetuadas. Adicionalmente foram também caraterizados e comparados outros subgrupos amostrais que se consideraram relevantes, com o objetivo de igualmente identificar a existência de padrões: amostras populacionais que avaliaram cada um dos quatro conjuntos de estados de saúde definidos; amostra com respostas válidas associadas ao método MACBETH com amostra com respostas válidas associadas ao método TTO; amostra populacional com exclusões MACBETH e com exclusões TTO; e amostras excluídas e incluídas para cada um dos métodos.

As escalas de valores médios obtidas de acordo com os dois métodos testados, bem como, a análise comparativa e a determinação da correlação entre estas são apresentadas na secção seguinte. Incluindo-se, também, nesta secção, a apresentação dos resultados obtidos em relação ao teste de ajuste da escala MACBETH à escala TTO, descrito na Secção 4.6.2.

A secção seguinte é dedicada à caraterização dos resultados obtidos no que diz respeito às preferências manifestadas pelos indivíduos em expressar julgamentos de valor em números ou palavras, tendo-se considerado de interesse, também aqui, analisar uma possível influência da ordem de apresentação das tarefas associadas a cada um dos métodos.

Por fim, na última secção deste capítulo, são apresentados os resultados obtidos de modo a dar resposta às questões levantadas relativamente à influência do grau de numeracia dos indivíduos. Nomeadamente é descrita a influência da numeracia nas preferências manifestadas, nas escalas de valores médios obtidas e na correlação entre estas e é, ainda, analisada a influência desta competência no número de exclusões identificadas na amostra.

5.1. Obtenção da Amostra Válida para análise

Na Figura 13 é apresentado um diagrama de fluxo da obtenção da amostra populacional considerada válida para análise. No total, 348 indivíduos, com idade superior a 18 anos e de

nacionalidade portuguesa, convidados por *email* a participar neste estudo, aceitaram responder ao questionário implementado, no entanto, apenas 243 responderam à sua totalidade, obtendo-se, assim, uma taxa de resposta de, aproximadamente, 70%.

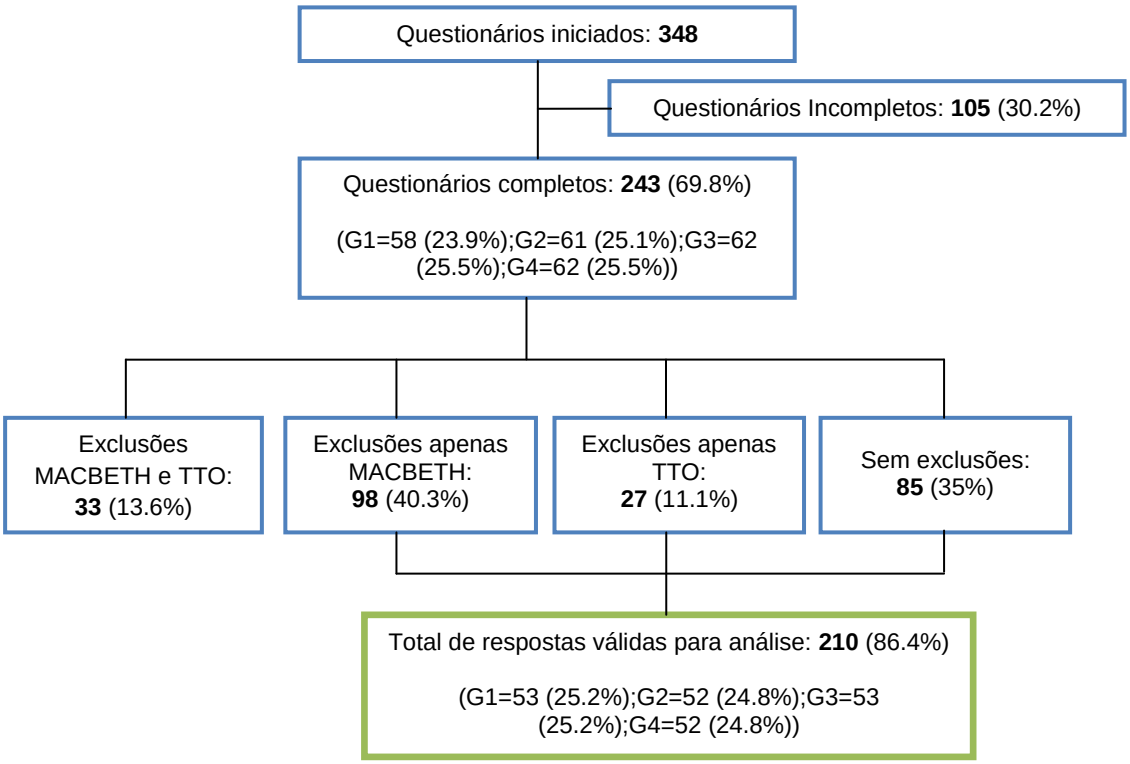

Figura 13. Diagrama de fluxo da obtenção da amostra populacional válida para análise (G1, G2, G3, G4 – Conjunto de questionários obtidos relativos ao grupo de estados de saúde 1, 2, 3 e 4, respetivamente).

Dado o modelo adotado para implementação do questionário na plataforma Qualtrics, em que foram introduzidos aspetos de aleatorização, nomeadamente na atribuição a cada indivíduo de um dos quatro conjuntos de estados de saúde anteriormente definidos para avaliação, é de interesse analisar que o número de respostas completas obtidas para cada um destes conjuntos foi, aproximadamente, igual.

Do total de 243 questionários completos obtidos, 33 foram excluídos da análise na sua totalidade, uma vez que, manifestaram pelo menos um dos critérios de exclusão (definidos na Secção 4.5.) em relação a ambos os métodos. Nos casos em que apenas num dos métodos foi identificado um ou mais dos problemas definidos como critérios de exclusão, excluíram-se apenas os julgamentos relativos a esse método. Assim, foram excluídos julgamentos apenas em relação ao método MACBETH em 98 questionários e apenas em relação ao método TTO em 27 questionários. Uma descrição do número de exclusões (totais e parciais) efetuadas para cada método, por critério de exclusão, é apresentada na Tabela 5. Nesta tabela são, ainda, reportados os resultados da análise dos dois critérios adicionais, relativos ao método TTO, definidos na Secção 4.5.

Obteve-se, então, um total de 210 questionários válidos para análise, entre os quais se obtiveram 112 questionários com julgamentos válidos de acordo com o método MACBETH e 183 questionários com julgamentos válidos de acordo com o método TTO. Verificou-se que o número total de questionários válidos obtidos se encontrava distribuído de forma igual entre os quatro conjuntos de

estados de saúde em avaliação, não se tendo verificado diferenças estatisticamente significativas em relação ao número de exclusões efetuadas, tendo em conta cada um dos critérios, nestes quatro conjuntos.

No total, apenas 85 (40.5%) dos questionários considerados válidos para análise incluíram julgamentos válidos associados tanto ao método MACBETH como ao método TTO.

Tabela 5. Descrição, por critério, do número total de julgamentos excluídos associados a cada método (N = 243 questionários completos).

Motivo de exclusão *	Julgamentos excluídos, n (%)
• MACBETH	**131 (53.9)**
1) Ordenação inconsistente	21 (8.6)
2) Todos os estados avaliados como piores do que a morte	4 (1.6)
3) Matriz de julgamentos inconsistente **	106 (43.6)
• TTO	**60 (24.7)**
1) Todos os estados avaliados como piores do que a morte	2 (0.8)
2) Mesmo valor atribuído a todos os estados	3 (1.2)
3) Para estados melhores do que a morte $t = 0$	17 (7)
4) Para estados piores do que a morte $t = 0$	49 (20.2)
• Critérios adicionais TTO	
Inconsistência lógica	126 (51.9%)
Inconsistência lógica grave	10 (4.1%)

* Em alguns questionários foi identificado mais do que um motivo de exclusão.
** No MACBETH a identificação deste motivo de exclusão apresenta uma ordem sequencial com os anteriores. Assim, respostas anteriormente excluídas de acordo com o critério 1) ou 2) não são analisadas de acordo o critério 3).

Influência da ordem de apresentação das tarefas

No desenvolvimento do questionário *online* implementado, conforme foi descrito anteriormente, a ordem de apresentação das tarefas de atribuição de julgamentos de acordo com o método baseado na abordagem MACBETH e de acordo com o método TTO foi aleatorizada, pelo que se considerou de interesse analisar a possível influência desta ordem em alguns dos dados recolhidos, nomeadamente no número de exclusões identificadas e no número de questionários incompletos obtido.

Verificou-se, assim, de acordo com a

Figura 14, que no que diz respeito ao total de questionários completos obtidos, a tarefa de atribuição de julgamentos semânticos, associada MACBETH, foi apresentada em primeiro lugar num maior número de questionários (55.1%). No entanto, relativamente ao número de exclusões totais

MACBETH e TTO, e relativamente ao número de exclusões apenas MACBETH e apenas TTO não se verificaram diferenças estatisticamente significativas a um nível de 5% entre os dois grupos.

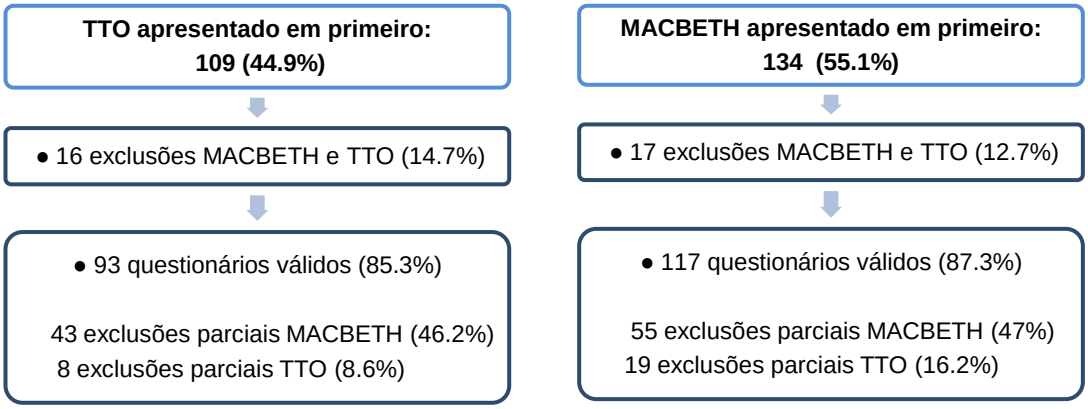

Figura 14. Descrição do número de exclusões obtidas tendo em conta a ordem de apresentação das tarefas de recolha de julgamentos de acordo com o método MACBETH e TTO.

Adicionalmente em relação aos questionários incompletos verificou-se que de entre o total de 105 questionários incompletos obtidos apenas em 50 questionários (47.6%) os indivíduos chegaram até ao módulo de atribuição de julgamentos de valor aos estados de saúde, sendo que entre estes, a tarefa de atribuição de julgamentos semânticos de acordo com o método MACBETH foi apresentada em primeiro lugar um maior número de vezes (62%).

Tendo em conta os dados apresentados na Figura 15, constatou-se uma proporção de indivíduos que concluíram a primeira tarefa apresentada, no caso de esta ser o método MACBETH estatisticamente superior aqueles que concluíram a primeira tarefa no caso de esta ser o TTO ($p = 0.02$). No entanto, as diferenças em relação ao número de indivíduos que concluíram a tarefa apresentada em segundo lugar não são estatisticamente significativas.

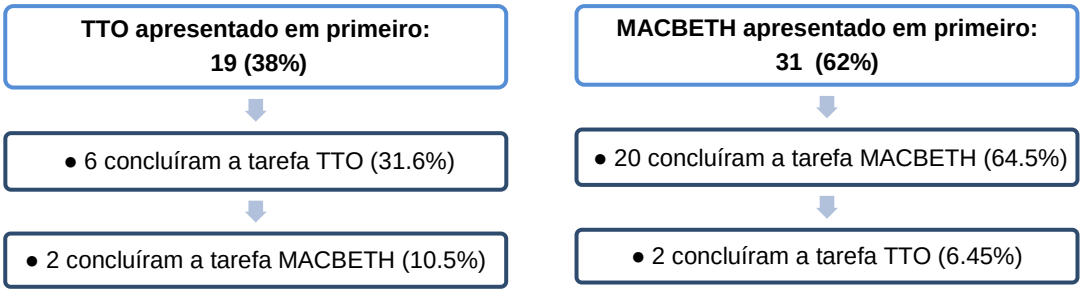

Figura 15. Descrição do número de tarefas concluídas, nos questionários incompletos, tendo em conta a ordem de apresentação das tarefas de recolha de julgamentos de acordo com o método MACBETH e TTO.

5.2. Caraterização da Amostra Populacional

5.2.1. Caraterização Sociodemográfica

As principais características sociodemográficas da amostra populacional deste estudo, antes e após exclusões, são apresentadas na Tabela 6. É possível observar-se que a amostra populacional considerada válida para análise (210 indivíduos) é maioritariamente do sexo feminino (63.8%) e apresenta uma idade média de, aproximadamente, 34 anos, sendo que a maior fração da população (33.3%) pertence à faixa etária 18-24 anos. Destaca-se, também, que relativamente ao nível de escolaridade da população a proporção de indivíduos a apresentar um nível superior ao Ensino Secundário (Bacharelato, Licenciatura, Mestrado ou Doutoramento) (78.3%) é bastante superior à proporção de indivíduos nas outras categorias. No que diz respeito à situação face ao trabalho os indivíduos encontram-se maioritariamente empregados (51.4%) ou numa situação de estudante (36.2%). É ainda de referir que, os indivíduos da amostra são maioritariamente solteiros (50%) ou casados/em união estável (48.1%) e, de modo geral, a maior fração da população não sente que a sua saúde e o acesso a cuidados de saúde/cuidados médicos sejam limitados pelo seu rendimento mensal líquido.

Relativamente à prevalência de doenças crónicas, a maior proporção de indivíduos não padece de nenhuma doença (79.1%). No que diz respeito aos doentes crónicos (18.6%), a maioria considera que a doença de que padecem não é limitativa da sua qualidade de vida nem impede a realização de atividades habituais adequadas à idade, podendo, no entanto, aumentar o risco de vir a sofrer de outras doenças.

No geral, as diferenças observadas nas características da amostra populacional antes e após exclusões são reduzidas, não sendo estatisticamente significativas a um nível de 5%.

Tabela 6. Caraterização da amostra populacional do estudo antes e após exclusões.

		Amostra total	Amostra após exclusões
N		**243**	**210**
Variáveis sociodemográficas		**n (%)**	**n (%)**
Sexo	**Feminino**	155 (64.0)	134 (63.8)
	Masculino	88 (36.0)	76 (36.2)
Idade	**Média ± DP**	34.4 ± 13.1	34.3 ± 13.1
	18-24	77 (31.7)	70 (33.3)
	25-30	46 (18.9)	37 (17.6)
	31-44	62 (25.5)	53 (25.2)
	> 44	58 (23.9)	50 (23.8)
Nível de escolaridade	**Inferior ao Ensino Secundário**	5 (2.1)	3 (1.4)
	Ensino Secundário	64 (26.3)	52 (24.8)
	Superior ao Ensino Secundário	174 (71.6)	155 (73.8)
Estado civil	**Solteiro**	124 (51.0)	105 (50.0)
	Casado/União estável	111 (45.7)	101 (48.1)

	Divorciado/Separado		6 (2.5)	3 (1.4)
	Viúvo		2 (0.8)	1 (0.5)
Situação profissional	**Estudante**		81 (33.3)	76 (36.2)
	Empregado		131 (53.9)	108 (51.4)
	Desempregado		9 (9.7)	7 (3.3)
	Reformado		10 (4.1)	9 (4.3)
	Doméstico		3 (1.2)	3 (1.4)
	Outra situação		9 (3.7)	7 (3.3)
Agregado familiar	**1-2 elementos**		69 (28.4)	56 (26.7)
	3-4 elementos		153 (63.0)	135 (64.3)
	5 ou mais elementos		21 (8.4)	19 (9.0)
Relativamente ao rendimento mensal líquido	**Não limita a minha saúde**	Discordo: Indiferente: Concordo: NS/NR:	68 (28.0) 24 (9.9) 149 (61.3) 2 (0.8)	54 (25.7) 21 (10.0) 133 (63.3) 2 (1.0)
	Não limita a minha qualidade de vida	Discordo: Indiferente: Concordo: NS/NR:	83 (34.2) 16 (6.6) 142 (58.4) 2 (0.8)	67 (31.9) 14 (6.7) 127 (60.5) 2 (1.0)
	Não limita o acesso que preciso a cuidados de saúde	Discordo: Indiferente: Concordo: NS/NR:	77 (31.7) 16 (6.6) 148 (60.9) 2 (0.8)	64 (30.5) 12 (5.7) 132 (62.9) 2 (1.0)
	Não limita os cuidados médicos que a minha saúde necessita	Discordo: Indiferente: Concordo: NS/NR:	78 (32.1) 19 (7.8) 144 (59.3) 2 (0.8)	65 (31) 15 (7.1) 128 (61.0) 2 (1.0)
	Deixa-me confortável que conseguirei lidar com eventuais problemas de saúde que surjam	Discordo: Indiferente: Concordo: NS/NR:	93 (38.3) 19 (7.8) 129 (53.1) 2 (0.8)	77 (36.7) 18 (8.6) 113 (53.8) 2 (1.0)
Doença crónica	**Sim**		47 (19.3)	39 (18.6)
	Não		189 (77.8)	166 (79.1)
	NS/NR		7 (2.9)	5 (2.4)
Relativamente à doença crónica	**Não limitam a minha qualidade de vida**	Discordo Indiferente: Concordo	22 (46.8) 0 25 (53.2)	18 (46.2) 0 21 (53.8)
	Não me impedem de realizar as atividades habituais que as pessoas da minha idade realizam	Discordo Indiferente: Concordo	12 (25.5) 5 (10.6) 30 (63.8)	10 (25.6) 5 (12.8) 24 (61.5)
	Não aumentam o risco de eu vir a sofrer de outras doenças	Discordo Indiferente: Concordo	27 (57.4)) 4 (8.5) 16 (34)	22 (56.4) 3 (7.7) 14 (35.9)
	Não diminuem os anos de vida que as pessoas da minha idade esperam viver	Discordo Indiferente: Concordo	16 (34.0) 7 (14.9) 24 (51.1)	13 (33.3) 7 (17.9) 19 (48.7)
Numeracia	**Média ± DP**		2.4 ± 0.8	2.4 ± 0.8
	0 respostas certas		7 (2.9)	6 (2.9)
	1 resposta certa		28 (11.5)	24 (11.4)
	2 respostas certas		67 (27.6)	55 (26.2)
	3 respostas certas		141 (58.0)	125 (59.5)

DP – Desvio-padrão

Adicionalmente foram, também, analisadas comparativamente as amostras populacionais que avaliaram cada um dos quatro conjuntos de estados de saúde (Tabela 16, Anexo B), tendo-se

identificado algumas diferenças estatisticamente significativas, nomeadamente, em relação ao número de indivíduos pertencentes à faixa etária 25-30 anos, sendo este número bastante reduzido para a amostra populacional do grupo 3 (3 indivíduos) em comparação com os restantes (14, 12, 8 indivíduos respetivamente nos grupos 1, 2 e 4). Verificaram-se, ainda, diferenças no que diz respeito à situação profissional na categoria 'Desempregado', que apresentou um número de indivíduos bastante superior para o grupo 4 (6 indivíduos) enquanto nos outros grupos apresentou nenhum ou apenas 1 indivíduo, e na categoria 'Outra situação' que para o grupo 2 e 4 apresenta não apresenta nenhum indivíduo, enquanto para os grupos 1 e 3 apresentou 3 e 4 indivíduos respetivamente. Por último, constatou-se, ainda, uma diferença relativamente ao agregado familiar, nomeadamente na categoria '5 ou mais elementos', que apresentou 9 indivíduos no grupo 1 e 6 no grupo 4 mas apenas 3 e 1 nos grupos 2 e 3 respetivamente. Analisando a Tabela 16 (Anexo B) é possível encontrar outras ligeiras variações, no entanto, estas não são estatisticamente significativas a um nível de 5%. De modo geral, após a interpretação das diferenças encontradas estas não foram consideradas pertinententes tendo em conta o propósito deste estudo.

No âmbito das caraterísticas sociodemográficas foram, ainda, realizadas comparações entre as amostras com respostas válidas associadas ao método MACBETH e com respostas válidas associadas ao método TTO, entre as amostras populacionais com exclusões MACBETH e com exclusões TTO e, por fim, entre as amostras excluídas e incluídas para cada um dos métodos. Para nenhum dos casos foram identificadas diferenças estatisticamente significativas a um nível de 5%, não sendo, assim, possível a definição de nenhum padrão associado às exclusões/inclusões para ambos os métodos.

5.2.2. Caraterização do Estado de Saúde Atual

Para além dos dados apresentados relativamente às características sociodemográficas da amostra, foram, também, recolhidos dados de caraterização do estado de saúde atual dos indivíduos, tendo-se para tal implementado o instrumento EQ-5D, cujos resultados se apresentam na Tabela 7.

De modo geral, os indivíduos da amostra apresentam um estado de saúde sem problemas (nível 1) em todas as dimensões do sistema descritivo considerado, exceto para as dimensões dor/mal-estar e ansiedade/depressão em que surgem alguns problemas moderados (nível 2), respetivamente, 21.4% e 24.3%.

O valor médio atribuído pelos indivíduos ao seu estado de saúde atual através do EQ-5D VAS é de 81.6 ± 13.2, o que está de acordo com o resultado obtido através do sistema classificação do EQ-5D refletindo um nível de saúde elevado da amostra.

Adicionalmente, a maior fração de indivíduos, 70.5%, considera o seu estado de saúde atual igual ao estado de saúde experienciado durante os últimos doze meses, enquanto 21.9% considera o estado atual melhor.

Não foram encontradas diferenças estatisticamente significativas a um nível de 5% entre a amostra populacional antes e após exclusões nem entre as restantes comparações entre subgrupos amostrais, anteriormente descritas.

Tabela 7. Distribuição das respostas dos indivíduos às dimensões do EQ-5D e ao EQ-5D VAS.

Estado de saúde atual – Amostra válida para análise (N=210)					
Dimensões EQ-5D					
Nível	**Mobilidade, n (%)**	**Cuidados Pessoais, n (%)**	**Atividades Habituais, n (%)**	**Dor/Mal-estar, n (%)**	**Ansiedade/Depressão, n (%)**
1	**207 (98.6)**	**209 (99.5)**	**205 (97.6)**	**164 (78.1)**	**155 (73.8)**
2	3 (1.4)	0	5 (2.4)	45 (21.4)	51 (24.3)
3	0	1 (0.5)	0	1 (0.5)	4 (1.9)
VAS					
Média ± DP			81.6 ± 13.2		
Intervalos de valores			[30,100]		

* Nível modal marcado a negrito
(1- sem problemas; 2 – problemas moderados; 3 – problemas severos)

5.2.3. Caraterização do Grau de Numeracia

De acordo com os dados relativos à numeracia apresentados na Tabela 6, verificou-se que a maior proporção da amostra populacional deste estudo respondeu corretamente às 3 questões de numeracia apresentadas (59.5%), sendo que apenas uma pequena proporção (2.9%) não respondeu acertadamente a nenhuma questão. Note-se que, não respostas foram consideradas como respostas erradas.

Não foram encontradas diferenças estatisticamente significativas a um nível de 5% entre a amostra populacional antes e após exclusões, nem entre os restantes subgrupos amostrais comparados.

5.3. Valoração dos Estados de Saúde

Nesta secção são descritos e analisados os resultados obtidos relativamente à valoração dos diferentes estados de saúde em avaliação recorrendo aos dois métodos testados: MACBETH e TTO.

Começam-se por descrever e comparar as escalas de valores médios obtidas, analisando-se as diferenças entre estas, nomeadamente o número de diferenças encontradas superior a 0.1 e a 0.05, conforme descrito na Secção 4.6.2. Prossegue-se com a determinação da correlação entre as escalas obtidas e, por fim, são apresentados os resultados obtidos em relação ao teste de ajuste da escala MACBETH à escala TTO, também descrito na Secção 4.6.2.

5.3.1. Comparação entre as Escalas

Na Tabela 8 são apresentados os valores médios obtidos para cada um dos estados de saúde, bem como o erro padrão associado à média, o número total de valorações obtidas, a percentagem de

valorações negativas e, ainda a diferença entre valores médios obtidos e o coeficiente de correlação obtido, para cada estado, entre os valores resultantes da abordagem MACBETH e do método TTO.

Tabela 8. Resultados obtidos relativamente à valoração do conjunto de 25 estados de saúde avaliados, recorrendo ao MACBETH e ao TTO.

	MACBETH			TTO				Correlação	
Estado	**n**	**Média ± SE**	**% val. neg.**	**n**	**Média ± SE**	**% val. neg.**	**Diferença**	**n**	**ρ**
11112	26	0.851 ± 0.026	0	48	0.811 ± 0.030	0	0.04	21	0.59
12111	28	0.860 ± 0.021	0	46	0.828 ± 0.021	0	0.032	22	0.11
11211	31	0.838 ± 0.024	0	42	0.800 ± 0.029	0	0.038	20	-0.07
21111	27	0.810 ± 0.021	0	47	0.777 ± 0.029	0	0.033	22	0.48
11121	27	0.854 ± 0.017	0	47	0.831 ± 0.025	0	0.023	22	0.74
11131	28	0.650 ± 0.037	0	46	0.555 ± 0.048	2.17	0.095	22	0.33
11113	31	0.576 ± 0.065	10.81	42	0.564 ± 0.064	7.14	0.012	20	0.54
11133	26	0.342 ± 0.066	15.38	48	0.404 ± 0.069	8.33	-0.062	21	0.72
22121	31	0.580 ± 0.050	3.23	42	0.675 ± 0.033	0	-0.095	20	0.28
11312	27	0.547 ± 0.032	0	47	0.533 ± 0.055	6.38	0.014	22	0.21
13311	26	0.434 ± 0.052	7.69	48	0.499 ± 0.043	2.08	-0.065	21	0.46
22122	27	0.506 ± 0.039	0	47	0.592 ± 0.037	0	-0.086	22	0.64
21323	28	0.351 ± 0.040	3.57	46	0.358 ± 0.045	2.17	-0.007	22	-0.06
32211	28	0.397 ± 0.050	3.57	46	0.263 ± 0.047	8.70	0.134	22	0.69
32313	31	0.148 ± 0.055	32.26	42	0.137 ± 0.075	21.43	0.011	20	0.64
22222	26	0.577 ± 0.040	0	48	0.458 ± 0.053	4.17	0.119	21	0.51
33232	31	0.039 ± 0.060	32.26	42	-0.214 ± 0.072	50	0.253	20	0.66
23232	27	0.276 ± 0.043	11.11	47	0.011 ± 0.071	25.53	0.265	22	0.34
13332	31	0.169 ± 0.058	29.03	42	0.008 ± 0.072	26.19	0.161	20	0.62
22233	28	0.259 ± 0.050	14.29	46	0.158 ± 0.066	17.39	0.101	22	0.37
32223	26	0.106 ± 0.059	34.62	48	0.061 ± 0.069	16.67	0.045	21	0.42
33321	26	0.145 ± 0.057	26.92	48	0.072 ± 0.075	20.83	0.073	21	0.54
23313	28	0.291 ± 0.039	3.57	46	0.177 ± 0.056	13.04	0.114	22	0.25
33323	27	0.046 ± 0.044	33.33	47	-0.269 ± 0.078	51.06	0.315	22	0.11
33333	112	-0.080 ± 0.020	48.24	183	-0.446 ± 0.035	63.93	0.366	85	0.41
						MAD:	**0.1015**		

SE – *Standard error*; **MAD –** *Mean Absolute Difference*; **ρ** – Correlação de Spearman

Tendo em conta a amostra populacional considerada válida para análise, constituída por um total de 210 indivíduos, em média cada estado de saúde foi avaliado recorrendo ao método MACBETH 28 vezes (intervalo entre 26 e 31) e recorrendo ao método TTO 47 vezes (intervalo entre 42 e 48), com exceção do estado 33333 que foi avaliado por todos os indivíduos apresentando, assim, 112 valorações recorrendo ao método MACBETH e 183 recorrendo ao método TTO.

O valor médio das valorações médias obtidas recorrendo ao método MACBETH é de 0.42 ± 0.29, sendo que os valores médios obtidos variam entre -0.080 (Estado 33333) e 0.860 (Estado 12111). Relativamente ao método TTO, o valor médio das valorações médias obtidas é de 0.35 ± 0.37, correspondendo o valor mínimo obtido a -0.446 (Estado 33333) e o valor máximo obtido a 0.831 (Estado 11121).

De modo geral, os valores médios obtidos não diferem acentuadamente entre os dois métodos. No entanto, verificam-se diferenças absolutas entre valores superiores a 0.1 para 9 estados (36%), nomeadamente, para os estados 33333, 33323, 23313, 22233, 13332, 23232, 33232, 22222, 32211, apresentados por ordem decrescente de diferença entre valores. Considerando uma diferença absoluta superior a 0.05 este valor aumenta para 15 estados (60%) em 25 estados de saúde totais. O valor médio da diferença absoluta (MAD) entre os valores médios obtidos através dos dois métodos é de 0.1015, sendo a diferença máxima obtida de 0.37 correspondente ao Estado 33333 e a diferença mínima obtida de 0.007 correspondente ao Estado 21323.

No que diz respeito à percentagem de valores negativos, verifica-se, de modo geral, que os estados considerados, por todos os participantes, melhores do que a morte no método TTO também o foram no método MACBETH. Nos casos em que tal não se verificou, a percentagem de valorações negativas obtidas, no método em que nem todos os indivíduos consideraram o estado melhor do que a morte, são baixas, revelando uma concordância entre os julgamentos efetuados. Ainda, relativamente à percentagem de valorações negativas, verifica-se que o método TTO apresenta maiores percentagens para estados de saúde mais negativos (com maiores limitações), enquanto para estados de saúde melhores o MACBETH apresenta percentagens maiores.

Na Figura 16 são representados visualmente, de forma simplificada, os valores médios atribuídos aos vários estados de saúde pelos dois métodos, anteriormente apresentados na Tabela 8, ordenados por ordem decrescente de acordo com a escala de valores obtida pelo MACBETH. Através desta análise visual é possível constatar que de modo geral, os valores médios obtidos recorrendo ao MACBETH são superiores aos obtidos recorrendo ao TTO, havendo ligeiras alterações desta ordem para estados ligeiros e moderados.

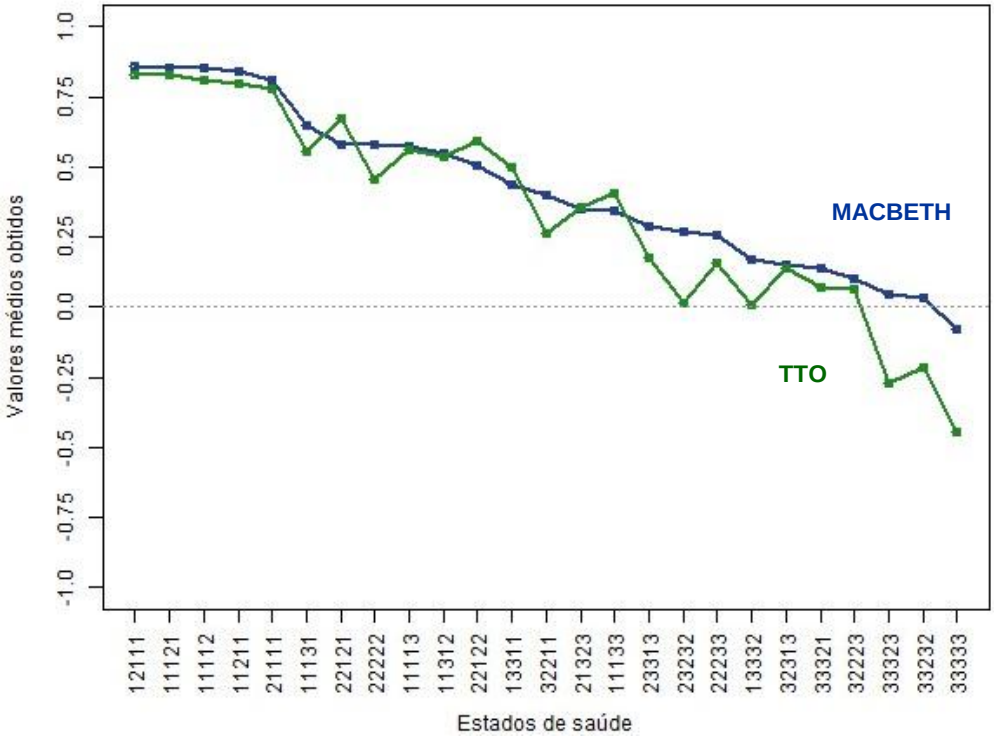

Figura 16. Representação do valor médio obtido para cada estado, ordenados de acordo com a escala de valores MACBETH. A cor verde representa-se os valores obtidos recorrendo ao TTO e a cor azul os valores obtidos recorrendo ao MACBETH.

5.3.2. Correlação entre Escala MACBETH e Escala TTO

Com o intuito de analisar, com maior detalhe, se os valores obtidos recorrendo ao método MACBETH se correlacionam com os valores obtidos recorrendo a um método de valoração tradicional, neste caso o TTO, foi determinado o coeficiente de correlação de Pearson, r . O valor obtido foi de, aproximadamente, 0.962 o que revela uma forte correlação positiva entre ambas as escalas de valores (p=1.6e-14).

Na Figura 17 é apresentada uma representação gráfica dos valores médios obtidos pelo método TTO em função dos valores médios obtidos pelo método MACBETH, evidenciando-se, a azul, a reta de regressão linear que descreve a relação entre as duas escalas. Verificando-se como esperado, de acordo com o valor de r obtido, uma boa aproximação dos pontos à reta de regressão linear. A Figura 17 permite, ainda, reforçar o resultado anteriormente apresentado de que, de modo geral, os valores obtidos recorrendo à escala MACBETH são superiores aos obtidos recorrendo ao TTO, uma vez que a maioria dos pontos encontra abaixo da reta $y = x$.

O coeficiente de determinação, r^2, obtido para este modelo é de 0.926, indicando uma relação linear forte entre as duas variáveis em que, aproximadamente, 93% da variabilidade encontrada para uma das escalas pode ser explicada pela outra.

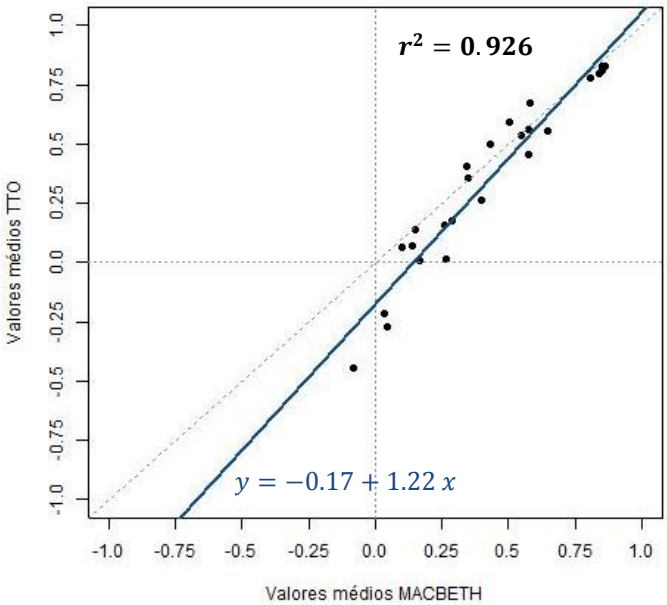

Figura 17. Valores médios atribuídos aos 25 estados de saúde pelo método TTO em função dos valores médios obtidos pelo método MACBETH.

Adicionalmente, à determinação do coeficiente de correlação de Pearson em relação às escalas de valores médios obtidas, determinou-se, também, a correlação entre as duas escalas obtidas para cada estado de saúde, tendo-se para tal, recorrido à correlação não-paramétrica de Spearman. Os diferentes valores de ρ obtidos foram apresentados anteriormente na Tabela 8, sendo que estes variam entre -0.006 (Estado 21323) e 0.74 (Estado 11121) e, de modo geral, verifica-se para os diferentes estados uma correlação moderada (0.4 < ρ ≤ 0.69). [106]

Na Figura 18 representam-se todos os pares de valores obtidos através do método MACBETH e do método TTO, verificando-se uma maior concentração de pontos no quadrante superior direito em torno da reta $y = x$, o que indica um certo grau de concordância entre os valores positivos obtidos pelas duas escalas. Destaca-se, ainda, a apresentação de alguns pontos sobre a reta $x = 0$, que dizem respeito a estados aos quais, recorrendo a método MACBETH, foi atribuído o valor 0, enquanto, recorrendo ao método TTO, foi atribuído um valor diferente de 0, positivo ou negativo. Embora a situação contrária também se verifique, esta não é tão evidente.

Da análise da Figura 18, nota-se, ainda, que para estados avaliados como negativos, de acordo com o método MACBETH, verifica-se uma concentração de pontos em redor da reta $y = 0$, aproximadamente igual na parte inferior e na parte superior da reta. Este ponto é ilustrativo da existência de vários estados avaliados como negativos pelo MACBETH e que no entanto de acordo com o método TTO foram avaliados como positivos. O contrário também se verifica, isto é, estados avaliados como positivos de acordo com o método MACBETH mas aos quais o valor atribuído pelo TTO é negativo, no entanto, os pontos que representam esta situação são em menor número (pontos no quadrante inferior direito).

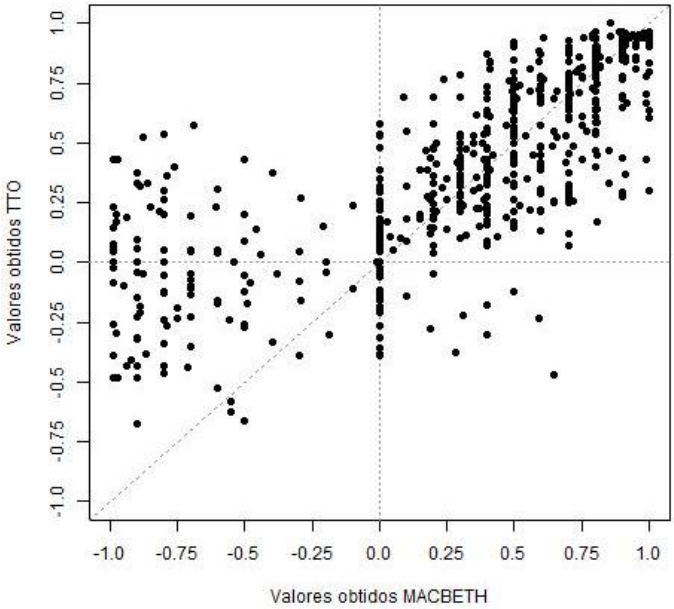

Figura 18. Valores atribuídos a todos os estados de saúde pelo método TTO em função dos valores atribuídos pelo método MACBETH.

5.3.3. Ajuste da Escala MACBETH à Escala TTO

Tendo em conta a funcionalidade do M-MACBETH, anteriormente referida na Secção 4.6.2., que permite ajustar a escala inicialmente proposta, dentro de determinados limites, mantendo a consistência da mesma com a matriz de julgamentos de valor, foi efetuado um teste de ajuste da escala MACBETH proposta à escala obtida para o mesmo indivíduo recorrendo ao método TTO. Como amostra de teste foi selecionada a população que avaliou o primeiro conjunto de estados de saúde (Grupo 1), tendo-se verificado que em nenhum caso a escala TTO foi consistente com a matriz de julgamentos MACBETH.

5.4. Caraterização das Preferências Manifestadas

Na Tabela 9 são apresentados os resultados obtidos em relação ao modo preferencial indicado pelos indivíduos da amostra para expressar julgamentos de valor. O modo dominante foi a atribuição de julgamentos semânticos recorrendo ao método MACBETH (47.6%), seguindo-se por uma elevada fração de indivíduos que consideraram indiferente a utilização de números ou categorias semânticas (30.5%).

Tabela 9. Caraterização do modo preferencial para expressão de julgamentos de valor manifestado pelos indivíduos da amostra.

Método preferencial	Amostra (N=210)
Julgamentos semânticos, n (%)	100 (47.6)
Julgamentos numéricos, n (%)	46 (21.9)
Indiferente, n (%)	64 (30.5)

De seguida, Tabela 10 são apresentados alguns dos principais comentários registados em relação a ambos os métodos implementados: MACBETH e TTO. Para ambos os métodos foram descritas diferentes vantagens e desvantagens referidas pelos indivíduos, bem como o número de vezes que foram indicadas.

Tabela 10. Principais comentários ao método baseado na abordagem MACBETH e ao método TTO.

Time Trade-Off			
Vantagens - (n)		**Desvantagens - (n)**	
Escala mais direta, com resposta mais precisa, menos sujeita a interpretações	6	Dificuldade em quantificar (traduzir em números) aspetos qualitativos e estabelecer uma equivalência com saúde perfeita	17
Mais fácil responder e ter quantificar uma escolha tendo como comparação a saúde perfeita	6	Dificuldades na visualização de horizontes temporais de 10 anos	3
Maior facilidade em manter e avaliar a coerência entre respostas dadas	3	Decisão difícil, sobretudo por se tratar de uma escala fina como de 0 a 10 anos com valores decimais	2
Mais simples e intuitivo	2	Tarefa mais complexa que exigiu mais tempo para compreensão	1
MACBETH			
Vantagens (n)		**Desvantagens (n)**	
Tarefa de comparação facilitada pela possível equiparação dos problemas	13	Resposta sujeita a diferentes interpretações, os julgamentos qualitativos podem não significar o mesmo para toda a gente	2
Maior facilidade na compreensão da tarefa pedida, mais percetível	10	Dificuldades em visualizar e imaginar os cenários descritos, sobretudo os estados mais negativos	1
Resposta mais direta, intuitiva e simples	6	Dificuldade em comparar estados muito semelhantes	1
Maior facilidade na atribuição de julgamentos baseados em categorias semântica	1		

Influência da ordem de apresentação das tarefas

Na Tabela 11 são descritas as preferências manifestadas em relação ao modo de expressão de julgamentos de valor tendo em conta a ordem de apresentação das duas tarefas de atribuição de julgamentos: MACBETH e TTO.

Tanto para a situação em que o método TTO é apresentado em primeiro lugar, como para a situação em que é o método MACBETH apresentado em primeiro lugar, o modo preferencial de expressão de julgamentos de valor são os julgamentos semânticos associados ao método MACBETH, 53.8% e 42.7% respetivamente, não sendo esta diferença estatisticamente significativa. No entanto, o mesmo não se verifica para a diferença entre o número de casos em que o modo preferencial de expressão de julgamentos de valor recai sobre os julgamentos numéricos, sendo que neste caso quando a tarefa MACBETH é apresentada em primeiro a preferência pelos julgamentos numéricos é superior ($p = 0.01$).

Aplicando o teste do χ^2 para independência à tabela de contingência, Tabela 11, obtém-se um valor $p = 0.04$, pelo que, a hipótese de que a preferência é independente da ordem de apresentação das tarefas deve ser rejeitada.

Tabela 11. Caraterização das preferências manifestadas de acordo com a ordem de apresentação das tarefas de atribuição de julgamentos.

	Tarefa TTO em primeiro: 93 (44.3%)	**Tarefa MACBETH em primeiro:** 117 (55.7%)
Julgamentos Semânticos, n (%)	50 (53.8%)	50 (42.7%)
Julgamentos numéricos, n (%)	13 (14.0%)	33 (28.2%)
Indiferente, n (%)	30 (32.2%)	34 (29.1%)

5.5. Caraterização da Influência do Grau de Numeracia

Nesta última secção, são apresentados e analisados os resultados obtidos referentes à influência do grau de numeracia da população, recolhidos com o intuito de dar resposta às questões levantadas na Secção 4. Assim, começa-se por descrever a influência do grau de numeracia nos resultados recolhidos relativamente ao modo preferencial de expressão de julgamentos de valor. De seguida, analisa-se a influência do grau de numeracia nas escalas de valores obtidas, para tal, determinam-se os coeficientes de correlação entre as escalas de valores médios MACBETH e TTO para amostras com diferentes graus de numeracia e representam-se visualmente as escalas de valores médios associadas a subgrupos amostrais com diferente grau de numeracia. Por fim, é, ainda, analisada a influência desta competência no número de exclusões identificadas na amostra.

Nos últimos dois pontos referidos, de modo a simplificar a apresentação de resultados, agruparam-se os indivíduos de acordo com duas categorias: “Elevada numeracia” (2-3 respostas certas) e “Baixa numeracia” (0-1 resposta certa), de acordo com Schwartz *et al.* (2004) [11]. Estas

categorias foram comparadas entre si no que diz respeito às variáveis sociodemográficas e não se identificaram diferenças estatisticamente significativas entre elas.

5.5.1. Influência do Grau de Numeracia nas Preferências

Na Tabela 12 são apresentados os resultados obtidos em relação ao modo preferencial de expressão de julgamentos de valor tendo em conta o grau de numeracia. Verifica-se que, para indivíduos com graus mais baixos de numeracia (apenas 1 ou nenhuma resposta certa) a preferência recai sobre a opção "Indiferente", enquanto, para indivíduos com maior grau de numeracia a escolha recai, maioritariamente, sobre a atribuição de julgamentos semânticos.

Tabela 12. Caraterização das preferências manifestadas relativamente ao modo de expressão de julgamentos de valor, de acordo com o grau de numeracia dos indivíduos.

	0 respostas certas (N=6)	**1 resposta certa (N=24)**	**2 respostas certas (N=55)**	**3 respostas certas (N=125)**
Julgamentos semânticos, n (%)	0	7 (29.2)	20 (36.4)	73 (58.4)
Julgamentos numéricos, n (%)	1 (16.7)	8 (33.3)	11 (20)	26 (20.8)
Indiferente, n (%)	5 (83.3)	9 (37.5)	24 (43.6)	26 (20.8)

Com o intuito de analisar a possível influência do grau de numeracia dos indivíduos na sua preferência por julgamentos de valor numéricos ou não numéricos foi aplicada uma regressão logística multinominal considerando como variável dependente a preferência manifestada pelos indivíduos (variável categórica com três níveis: "Julgamentos semânticos", "Julgamentos numéricos" e "Indiferente") e como variável independente o grau de numeracia (variável numérica com valores entre 0 e 3).

Usando como referência a categoria "Indiferente" verificou-se que a variável independente é estatisticamente significativa na distinção entre as categorias "Indiferente" e "Julgamentos semânticos" (OR = 2.50 (95% CI: 1.63-3.86)), registando-se uma associação positiva entre as variáveis. Assim, os indivíduos com maior grau de numeracia têm maior probabilidade de se encontrar no grupo que prefere expressar julgamentos de valor em palavras do que no grupo de indivíduos indiferentes. No entanto, a variável independente não é estatisticamente significativa na distinção entre as categorias "Indiferente" e "Julgamentos numéricos" (OR = 1.30 (95% CI: 0.87-2.14))

Por outro lado, usando como referência "Julgamentos numéricos" verificou-se, ainda, que o grau de numeracia é estatisticamente significativo na distinção entre a categoria de referência e a categoria "Julgamentos semânticos" (OR = 1.83 (95% CI: 1.13-3.00), sendo que, um maior grau de numeracia aumenta a probabilidade de preferência por expressão de julgamentos de valor através de categorias semânticas, em relação à probabilidade de expressar julgamentos de valor em números.

Para ambos os modelos apresentados, a variável independente explica 4.4% (McFadden R^2) da variabilidade observada na variável dependente, que embora seja um valor baixo é estatisticamente significativo.

5.5.2. Influência do Grau de Numeracia na Valoração de Estados de Saúde

Na Tabela 13 são apresentados os coeficientes de correlação de Spearman (ρ) entre as escalas de valores médios MACBETH e TTO obtidas, de acordo com o grau de numeracia dos indivíduos. Adicionalmente são, ainda, apresentados os coeficientes de determinação (ρ^2) que explicam a variabilidade de uma escala de valores em função da outra.

Tabela 13. Correlação de Spearman entre escalas de valores obtidas recorrendo ao método MACBETH e TTO.

	ρ (valor p)	**ρ^2**
0 respostas certas	0.378 (0.11)	0.143
1 resposta certa	0.905 (2.01e-06)	0.819
2 respostas certas	0.895 (2.14e-06)	0.801
3 respostas certas	0.965 (8.98e-07)	0.931

Verifica-se que, quanto maior o grau de numeracia dos indivíduos maior a correlação entre as escalas de valores médios obtidas por ambos os métodos, registando-se um coeficiente de correlação máximo (0.965), muito próximo de 1, para o grau mais elevado de numeracia.

Na Figura 19 e na Figura 20 são representadas ambas as escalas de valores médios obtidas pelos métodos TTO e MACBETH, respetivamente, para a amostra classificada na categoria “Elevada numeracia” e para amostra classificada na categoria “Baixa numeracia”. É de interesse notar, que de modo geral para ambos os métodos se verifica uma aproximação dos valores médios obtidos pelas duas categorias, no entanto, verificam-se algumas discrepâncias acentuadas, nomeadamente para o estado 11113 no método MACBETH. A média das diferenças absolutas entre as escalas de valores médios TTO obtidas por ambas as amostras, foi de 0.120 e a média das diferenças absolutas entre as escalas de valores MACBETH foi de 0.128, pelo que, em primeira análise, se verifica que relativamente às diferenças observadas entre amostras o grau de numeracia afetou de igual modo os dois métodos testados.

Verificou-se, ainda, que o método MACBETH apresentou um maior número de casos em que de acordo com uma amostra os estados são classificados, em média, como melhores do que a morte e de acordo com a outra são classificados, em média, como piores do que a morte.

Em 13 estados de saúde de um total de 25, no método TTO a amostra com baixa numeracia apresenta valores médios atribuídos inferiores. Por outro lado, no método MACBETH em 18 estados de saúde de entre 25, a amostra com baixa numeracia apresenta valores médios atribuídos inferiores.

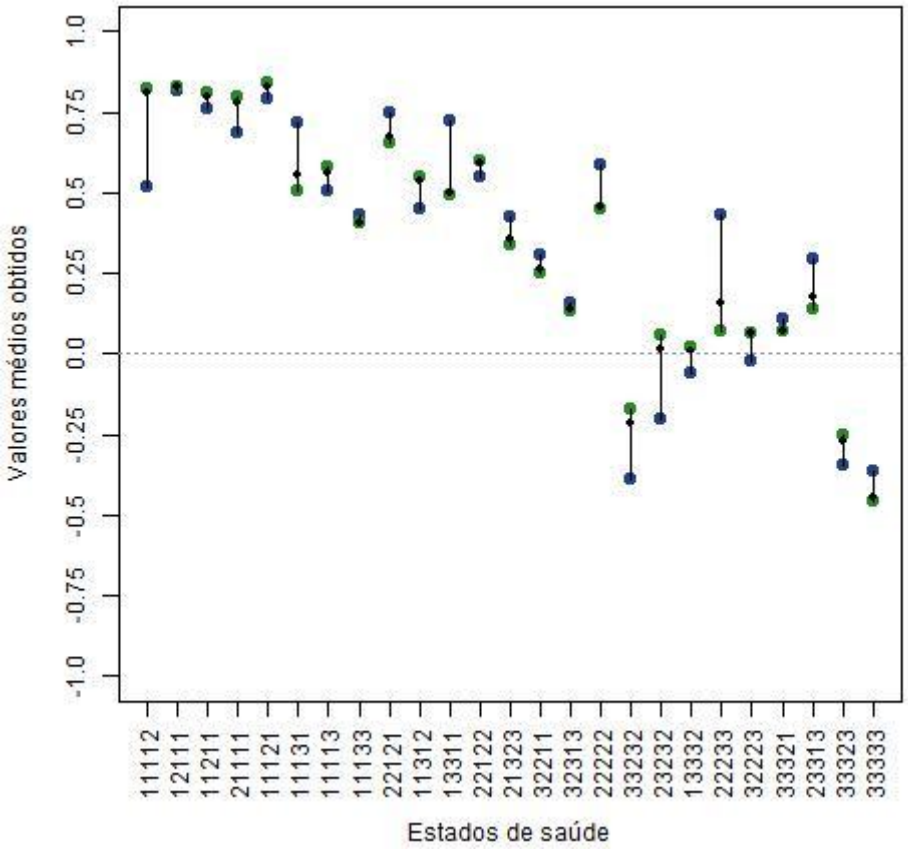

Figura 19. Representação dos valores obtidos recorrendo ao método TTO para a amostra total (ponto preto), para a amostra com elevada numeracia (ponto verde) e para a amostra com baixa numeracia (ponto azul).

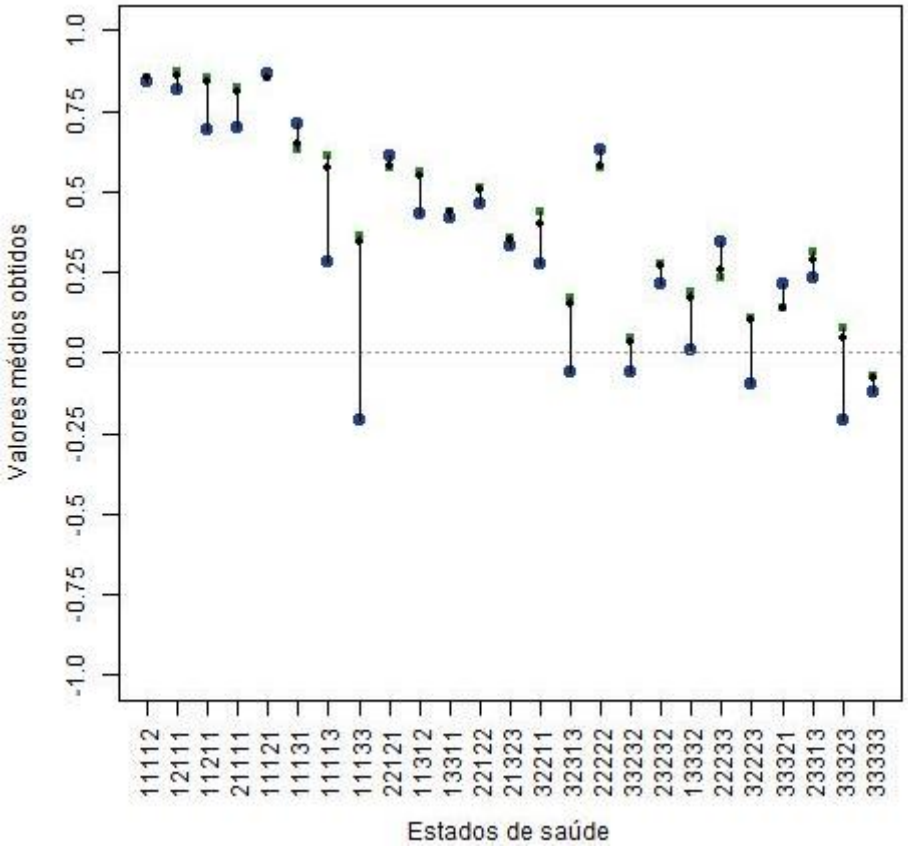

Figura 20. Representação dos valores obtidos recorrendo ao método MACBETH para a amostra total (ponto preto), para a amostra com elevada numeracia (ponto verde) e para a amostra com baixa numeracia (ponto azul).

5.5.3. Influência do Grau de Numeracia nas Exclusões

Na Tabela 14 são apresentados o número de exclusões identificadas para as amostras com baixo grau de numeracia e com elevado grau de numeracia, tendo em conta a população inicial do estudo (243 indivíduos). Não se obtiveram diferenças estatisticamente significativas no número de exclusões MACBETH e TTO, bem como no número de exclusões apenas TTO ou apenas MACBETH entre ambas as amostras, no entanto, ao analisar o critério adicional TTO o mesmo não se verifica (p=0.04 e p=0.02, respetivamente para as inconsistências lógicas e para as inconsistências lógicas graves). Assim, verifica-se que indivíduos com baixo grau de numeracia apresentam um maior número de inconsistências lógicas associadas ao método TTO.

Tabela 14. Descrição do número de exclusões identificadas para a amostra com "Baixa numeracia" e "Elevada numeracia".

	Baixa numeracia (N=35)	**Elevada numeracia (N=208)**
Exclusões MACBETH e TTO	5 (14.3%)	28 (10.0%)
Exclusões apenas TTO	1 (2.9%)	26 (12.5%)
Exclusões apenas MACBETH	16 (45.7%)	82 (39.4%)
Critério adicional TTO*	(N=30)	(N=180)
- Inconsistência lógica	18 (60%)	72 (40%)
- Inconsistência lógica grave	3 (10%)	4 (2.2%)
* O critério adicional TTO foi analisado apenas para a amostra populacional válida (210 indivíduos)		

6. Discussão

A aplicação da metodologia proposta resultou no desenvolvimento e teste de um novo protocolo não numérico de avaliação de estados de saúde conjuntamente com o método convencional TTO e, ainda, na determinação de preferências em relação a ambos e inferências relativamente à influência do grau de numeracia. Os resultados obtidos foram apresentados no capítulo anterior, sendo o propósito deste capítulo a realização de uma análise crítica dos resultados mais importantes dando resposta aos objetivos definidos e às questões anteriormente levantadas neste estudo. Em particular, comparam-se os dois métodos de avaliação de estados de saúde implementados, discutindo-se as escalas de valor obtidas, bem como, as preferências manifestadas pelos indivíduos e é dada resposta às questões relativas à numeracia enunciadas na Secção 4. Adicionalmente comparam-se as escalas de valores obtidos com a escala de valores definida para os estados de EQ-5D em Portugal e, por fim, referem-se algumas das limitações identificadas na metodologia proposta.

6.1. Comparação entre MACBETH e TTO

A primeira motivação para o desenvolvimento deste estudo consistiu em propor um novo método não numérico de avaliação de estados de saúde e testá-lo conjuntamente com um método numérico, tendo-se para tal, selecionado o método *Time Trade-Off*. Deste modo, possibilitou-se a comparação entre ambos os métodos em termos do número de exclusões e do número de desistências, da escala de valores obtidas, da preferência manifestada relativamente a estes e, ainda, do tempo de resposta.

Relativamente ao número de exclusões identificadas em cada um dos métodos (Tabela 5), verificou-se, em primeira análise, um número de exclusões associadas ao MACBETH superior, no entanto, é necessário ter em conta que os critérios de exclusão associados a cada método diferem entre si, estando de acordo com o protocolo seguido. Destaca-se, assim, a observação de que, enquanto no MACBETH se excluíram 21 respostas (8.6%) em que a tarefa de ordenação dos estados originou inconsistências lógicas, embora este não tenha sido um critério de exclusão explícito considerado no TTO, foi um critério adicional analisado, tendo-se constatado um total de 126 respostas (51.9%) com inconsistências lógicas na ordenação dos estados de saúde, das quais 10 (4.1%) foram classificadas como inconsistências lógicas graves, sendo este valor bastante superior ao obtido pelo MACBETH.

Ainda relativamente às exclusões identificadas é relevante referir que a maior proporção de exclusões associadas ao MACBETH foi justificada pela obtenção de inconsistências na matriz de julgamentos (43.6%), pelo que a incorporação no protocolo de questionário proposto de uma avaliação de inconsistências simultaneamente à recolha de julgamentos poderia evitar este tipo de exclusões, diminuindo-se, assim, de forma significativa o número total de exclusões associadas ao método MACBETH.

Outro aspeto a ter em consideração é o fato de que, embora, a ordem de apresentação das tarefas de recolha de julgamentos associada a cada um dos métodos não conduza a uma diferença estatisticamente significativa no número de exclusões identificadas, o mesmo não se verifica

relativamente ao número de desistências obtidas. Assim, verificou-se que, de entre os indivíduos que não completaram o questionário na sua totalidade, aqueles a quem foi apresentada em primeiro lugar a tarefa de atribuição de julgamentos associada ao MACBETH avançaram mais no preenchimento do questionário, completando a totalidade da tarefa MACBETH em 64.5% dos casos. Enquanto dos indivíduos a quem foi apresentada a tarefa TTO em primeiro apenas 31.6% concluíram esta tarefa.

Tendo em conta o desenho deste estudo, não é possível estabelecer objetivamente qual dos métodos revelou uma maior facilidade de compreensão das tarefas pedidas, no entanto, o resultado anteriormente referido pode ser indicativo de uma maior facilidade de compreensão e de resposta às questões associadas ao método MACBETH. Sendo esta conclusão reforçada pela constatação de uma maior preferência manifestada pelos indivíduos pelo modo de expressar julgamentos de valor associado a este método (julgamentos semânticos de diferença de atratividade entre opções) e pelos principais comentários identificados na Tabela 10, bem como, um menor tempo médio de resposta (aproximadamente 4 minutos *vs.* aproximadamente 5 minutos e meio obtidos para o método TTO). Em relação a este último ponto importa evidenciar que, tendo em conta o fato do questionário ser *online* e de preenchimento autónomo não é possível assegurar que não tenham existido interrupções nas tarefas de recolha de julgamentos MACBETH e TTO e, como tal, os tempos de resposta obtidos podem não ser precisos, uma vez que o tempo contabilizado corresponde ao tempo de permanência na página associada à recolha de julgamentos para cada um dos métodos. Os resultados até agora discutidos evidenciam, assim, a potencialidade de uso do novo método proposto como alternativa aos métodos convencionais de avaliação de estados de saúde.

No que diz respeito à análise das escalas de valores médios obtidas recorrendo ao novo método proposto e ao método TTO (Tabela 8), é importante referir que, embora o número de valorações obtidas seja reduzido (em médio 28 valorações MACBETH e 47 valorações TTO), de acordo com Bansback *et. al.* (2012) [92], 25-50 valorações para cada estado são suficientes para avaliar possíveis heterogeneidades nas preferências, sendo que os números obtidos neste estudo se encontram nesse intervalo. Importa, também, realçar que, no caso, de se ter optado por considerar como amostra válida para análise apenas os questionários com atribuições de julgamentos válidas para os dois métodos (85 questionários) o ponto anterior não se verificaria.

Analisando a associação entre as escalas de valores médios obtidas constatou-se que estas se correlacionam fortemente (r=0.962) e, que, de modo geral, as diferenças entre os valores médios de ambas as escalas são reduzidas tendo-se obtida uma diferença média absoluta de 0.1. Note-se que, enquadrando este valor relativamente à escala TTO uma diferença de 0.1 equivale a uma variação de um ano, para estados melhores do que a morte.

Para estados de saúde severos registam-se as maiores diferenças absolutas obtidas entre as escalas, por exemplo, a maior diferença entre valores médios apresentou-se para o estado 33333 que na escala MACBETH obteve um valor de -0.080 enquanto na escala TTO o mesmo foi de -0.446, o que pode ser justificado tendo em conta a transformação aplicada na conversão das escalas negativas numa escala entre -1 e 0, pois embora se tenha aplicado a mesma transformação a ambas as escalas os limites das escalas originais obtidas diferiam entre si, sendo o limite inferior da escala original TTO de -99 enquanto na escala MACBETH esse valor não é determinado.

Apesar da forte correlação entre as escalas de valores médios obtidas verificou-se que, ao tentar ajustar a escala de valores MACBETH de base, obtida para um determinado indivíduo, aos valores obtidos pelo mesmo recorrendo ao método TTO, a matriz de julgamentos MACBETH não é consistentes com a escala TTO.

É importante evidenciar que os valores TTO obtidos não podem ser considerados como o "*gold standard"* com os quais os valores obtidos pelo MACBETH devem ser comparados, uma vez que ambos resultam de tarefas diferentes, baseadas em princípios diferentes. Os valores obtidos pelo TTO proporcionam, no entanto, uma base de comparação que possibilita a contextualização dos resultados obtidos, bem como, a discussão das implicações resultantes da utilização de cada um dos métodos. Assim, de modo geral, constatou-se que os valores obtidos pelo MACBETH são superiores aos valores obtidos recorrendo ao método TTO (valor médio da escala MACBETH é de 0.42 e valor médio da escala TTO é de 0.35), constando-se, ainda que relativamente ao método MACBETH o número de valorações negativas obtidas para estados com maiores limitações é inferior. Este fato pode ser indicativo de uma maior tendência para que, de acordo com o protocolo MACBETH, numa maior percentagem de vezes os estados sejam classificados, por comparação direta, como melhores do que a morte, enquanto relativamente ao método TTO esta comparação não é direta resultando da idealização de uma situação em que se viveria 10 anos num dado estado de saúde e são apresentadas ao indivíduo duas hipóteses: viver 10 anos ou morrer imediatamente, sendo que estas abordagens podem ter um efeito psicológico diferente.

6.2. Considerações relativamente à Numeracia

No que diz respeito às questões levantadas em relação à influência do grau de numeracia da população tanto nas preferências manifestadas pelos indivíduos como nas próprias escalas de valores obtidas, verificou-se que para as diferentes hipóteses em estudo esta influência foi válida.

Analisando os resultados obtidos determinou-se que, a escolha do modo preferencial de expressão de julgamentos de valor, em números ou palavras, é influenciada pelo grau de numeracia da amostra, sendo que, com o aumento do grau de numeracia aumenta a preferência pela expressão de julgamentos de valor em palavras, isto é, a preferência pelo novo método proposto baseado na abordagem MACBETH. Tendo em conta as referências bibliográficas consultadas, nomeadamente o estudo Fasolo e Bana e Costa (2014) [12], seria esperado que indivíduos com maior grau de numeracia apresentassem preferência pela expressão de julgamentos de valor em números, no entanto, numa primeira análise, o estudo realizado constata que indivíduos com maior grau de numeracia apresentam uma maior preferência pela expressão de julgamentos em palavras. No entanto seria de interesse avaliar também a fluência dos indivíduos e analisar a influência adicional desta competência.

Outra constatação relativamente à numeracia foi a influência desta nas escalas de valores obtidas, bem como, na correlação obtida entre escalas derivadas de ambos os métodos para amostras com diferentes graus de numeracia. Comprovou-se, pois, que tal como esperado de acordo com Woloshin *et al.* (2001) [10] e Schwartz *et al.* (2004) [11], as utilidades médias obtidas diferem de

acordo com o grau de numeracia da amostra populacional, não se identificando no entanto diferenças significativas nesta variação entre ambos os métodos, e relativamente à correlação entre escalas de valores obtidas pelos dois métodos para amostras populações com diferentes graus de numeracia, confirmou-se o resultado indicado em Schwartz *et al.* (2004) [11], isto é, maior correlação entre escalas de valores obtidas pelos dois métodos para indivíduos com maior numeracia.

Adicionalmente analisaram-se, ainda, as exclusões identificadas para cada método de acordo com o grau de numeracia da amostra, tendo-se concluído que, embora não se tenham verificado diferenças estatisticamente significativas entre o número de exclusões objetivas para uma amostra com elevado grau de numeracia (2-3 respostas certas) e outra com baixo grau de numeracia (0-1 respostas certas), registou-se um número de inconsistências lógicas associadas ao TTO superior para uma amostra com baixa numeracia.

6.3. Comparação com Escala de Valores para o EQ-5D Portugal

Nesta secção pretende-se analisar a associação entre as escalas de valores médios obtidas neste estudo, recorrendo ao método MACBETH e ao método TTO, com a escala de valores médios de estados do EQ-5D atualmente definida em Portugal.

A metodologia desenvolvida, nomeadamente o questionário de avaliação de estados de saúde, teve por base uma implementação Web, diferindo, por isso, à partida da metodologia aplicada para obtenção da escala de valores atualmente existente para os estados de saúde descritos de acordo com o EQ-5D em Portugal. [76] Adicionalmente, o questionário desenvolvido pretendeu ainda incorporar outros aspetos que não apenas a implementação do método convencional TTO pelo que de modo a simplificar e reduzir a carga temporal do mesmo foi necessário adaptar o protocolo do método TTO (conforme descrito anteriormente na Secção 4.3.4.2) sendo neste ponto introduzida, mais uma vez, margem para alguns desvios às escala de valores médios TTO obtidas recorrendo aos protocolos convencionalmente estabelecidos. Diferentes implementações do protocolo associado ao método TTO originam diferentes escalas de valores. [107] [108]

Comparando visualmente a população da amostra válida deste estudo com a população considerada no estudo Ferreira *et al.* (2014) [76] identificaram-se algumas diferenças, destacando-se a idade média da população, que no presente estudo é bastante inferior (34.3 *vs.* 49.5), bem como o maior nível de escolaridade da população, sendo que, 73.8% dos indivíduos apresentam um nível de escolaridade superior ao Ensino Secundário (*vs.* 36.9% no estudo Ferreira *et al.* (2014)) e, ainda, a observação de que 36.2% ainda se encontram numa situação de estudante (*vs.* 9.6%). Destaca-se, também, o fato de a população deste estudo apresentar uma menor proporção de indivíduos com doenças crónicas (apenas 18.6% *vs.* 44.4%).

Ainda assim, os resultados obtidos pelos dois estudos apresentam uma forte correlação de acordo com a Tabela 15.

Tabela 15. Coeficiente de correlação de Pearson entre as escalas de valores médios obtidas no presente estudo (TTO e MACBETH) e a escala definida para a população portuguesa.

	TTO *vs.* TTO PT	**MACBETH *vs.* TTO PT**
Correlação de Pearson (r)	0.92	0.96
Valor p	6.4e-11	1.8e-14

6.4. Limitações da metodologia proposta

Nesta secção são descritas algumas das limitações da metodologia proposta que foram sendo identificadas no desenvolver deste estudo, sendo que alguns dos pontos de seguida referidos constituem, também, indicações para desenvolvimento futuro da metodologia proposta.

No que diz respeito ao novo protocolo não numérico de avaliação de estados de saúde baseado na abordagem MACBETH identificam-se como principais limitações o fato de a metodologia proposta não incluir uma avaliação da consistência dos julgamentos semânticos efetuados simultânea com a recolha dos mesmos, o que diminuiria acentuadamente o número de exclusões associadas e esta metodologia. Adicionalmente, evidencia-se, também, o fato de que o protocolo implementado não considerou a possibilidade de ajuste da escala proposta pelo *software* M-MACBETH, o que permitiria assegurar uma maior concordância entre a escala obtida e as opiniões manifestadas pelos indivíduos.

Considerando o protocolo seguido para implementação do método TTO, destaca-se que se recorreu a uma adaptação do mesmo e não ao protocolo convencionalmente seguido nos estudos mais recentes da área, pelo que a introdução destas variações metodológicas pode originar diferenças nos resultados obtidos, contudo seguindo-se a abordagem convencional seria necessário ter em atenção a gestão da carga temporal do questionário. Poderia ser possível diminuir o número de estados de saúde avaliados por cada individuo, criando-se um maior número de grupos o que por outro lado implicaria uma amostra de indivíduos com maior dimensão.

Relativamente ao desenvolvimento do questionário para recolha de julgamentos numéricos e não numéricos, bem como, de dados relativos às preferências e à numeracia, optou-se com o intuito de simplificar e diminuir a carga temporal do mesmo em não incluir um teste de literacia, contudo em desenvolvimentos futuros considera-se que o estudo da influência desta competência é um fator de interesse, nomeadamente a determinação da influência do grau de literacia nas preferências manifestadas relativamente ao modo de expressão de julgamentos de valor.

Por último, foram identificadas limitações associadas à estratégia de amostragem seguida, nomeadamente, amostragem não probabilística por redes, uma vez que este tipo de estratégia não assegura a representatividade da amostra em relação à população de interesse, neste caso a população portuguesa.

7. Conclusão

Nos últimos anos a Avaliação de Tecnologias em Saúde tem registado um desenvolvimento crescente como resultado da, também, crescente necessidade de uma alocação informada dos limitados recursos disponíveis. Dentro da ATS as avaliações económicas têm vindo a adquirir cada vez maior importância, nomeadamente as análises custo-utilidade que, recorrem à determinação de medidas únicas, como os QALYs, para quantificar os benefícios em saúde obtidos.

A abordagem dos QALYs é caraterizada por incorporar num único índice alterações na quantidade e na qualidade de vida sendo que, a QVRS associada a diferentes estados de saúde requer a modelação de um valor único para cada estado tendo como base as preferências dos indivíduos. Os diferentes métodos de elicitação de preferências em saúde são, hoje em dia, amplamente debatidos e estudados, tendo em conta a relevância crescente desta área, contudo não existe um consenso em relação a qual o método mais adequado, verificando-se que a cada método são associadas vantagens e desvantagens distintas.

As principais desvantagens associadas aos métodos convencionais, nomeadamente ao SG e ao TTO, estão associados a uma elevada exigência cognitiva e dificuldade de administração dos protocolos que lhes estão associados, sendo que estudos recentes sugerem, ainda, que os valores obtidos são influenciados por competências dos indivíduos, tais como a numeracia.

Neste contexto, o âmbito deste estudo consistiu no desenvolvimento de um novo protocolo, não numérico, de avaliação de estados de saúde tendo como base a abordagem MACBETH (*Measuring Attractiveness by a Categorical Based Evaluation Technique*) e no teste do mesmo conjuntamente com um método numérico convencional, neste caso o TTO, para obtenção de escalas de valores médios de utilidades de saúde associadas a diferentes estados descritos de acordo com o EQ-5D. Adicionalmente foram, também, determinadas as preferências dos indivíduos em relação a ambos os métodos e inferiu-se em relação à influência da numeracia nas preferências manifestadas e nas escalas de valores obtidas. Para implementação dos protocolos de recolha de julgamentos de valor numéricos e não numéricos, associados ao TTO e ao MACBETH respetivamente, e para recolha de dados relativos às preferências e ao grau de numeracia propôs-se o desenvolvimento de um questionário *online* a distribuir a uma amostra da população geral portuguesa.

Os resultados obtidos neste estudo determinaram que o novo método proposto baseado na abordagem MACBETH permite a valoração de diferentes estados de saúde, obtendo-se resultados lógicos e consistentes com aqueles obtidos recorrendo ao método TTO, sendo que as duas escalas obtidas apresentam uma elevada correlação entre si. Adicionalmente verificou-se, ainda, uma maior proporção de indivíduos que manifestaram preferência por expressar julgamentos de valor recorrendo a categorias semânticas de diferenças de atratividade, características da abordagem MACBETH, o que reforçou o interesse no estudo e desenvolvimento do método proposto. Inclusivamente, indivíduos com elevado grau de numeracia manifestaram preferência pela atribuição de julgamentos semânticos.

Evidenciou-se, assim, a potencialidade de uso do novo método proposto como alternativa aos métodos convencionais de avaliação de estados de saúde. Destaca-se, que o método baseado na

abordagem MACBETH, além de ser preferido pelos indivíduos, permite superar algumas das limitações associadas ao TTO, não só por ser cognitivamente mais simples e intuitivo, não requerendo julgamentos numéricos, mas também porque permite a valoração de estados de saúde considerados melhores e piores do que a morte recorrendo a um procedimento único, em que não é necessária uma etapa prévia de identificação relativamente a cada um dos estados se este é considerado melhor, pior ou igual à morte, tal como se verifica no protocolo TTO. No protocolo MACBETH é apenas requerida uma tarefa prévia de ordenação dos estados, sendo que esta já é atualmente apresentada nas valorações recorrendo ao TTO como exercício de “aquecimento”.

Contudo, ressalva-se a necessidade de desenvolvimentos futuros do método nomeadamente no que diz respeito à recolha de julgamentos simultânea com a avaliação da consistência dos mesmos de modo a diminuir o número de exclusões. Adicionalmente, é, ainda, relevante a replicação do estudo a uma amostra maior da população geral portuguesa, bem como, em contextos de doenças específicas e para populações específicas.

Referências

[1] "OECD." [Online]. Available: http://www.oecd.org/statistics/.

[2] S. Robinson, H. Dickinson, I. Williams, T. Freeman, B. Rumbold, K. Spence, H. Services, and M. Centre, "Setting priorities in health," *Heal. Serv. Manag. Centre, Univ. Birmingham Nuff. Trust*, no. September, 2011.

[3] V. Diaby, K. Campbell, and R. Goeree, "Multi-criteria decision analysis (MCDA) in health care: A bibliometric analysis," *Oper. Res. Heal. Care*, vol. 2, no. 1–2, pp. 20–24, Mar. 2013.

[4] V. Silva and Ó. Lourenço, "Avaliação Económica de programas de saúde - Essencial sobre conceitos, metodologia, dificuldades e oportunidades," *Rev. Port. Clínica Geral*, pp. 729–752, 2008.

[5] R. Taylor and R. Taylor, "What is health technology assessment?," *Hayward Med. Commun.*, no. 4, pp. 1–6, 2009.

[6] T. Walley, "Health technology assessment in England: assessment and appraisal," *MJA*, vol. 187, no. 5, pp. 283–285, 2007.

[7] M. Johannesson, "Economic evaluation of health care and policymaking," *Health Policy (New. York).*, vol. 33, pp. 179–190, 1995.

[8] S. J. Whitehead and S. Ali, "Health outcomes in economic evaluation: the QALY and utilities.," *Br. Med. Bull.*, vol. 96, pp. 5–21, Jan. 2010.

[9] E. Stamuli, "Health outcomes in economic evaluation: who should value health?," *Br. Med. Bull.*, vol. 97, pp. 197–210, Jan. 2011.

[10] S. Woloshin, L. M. Schwartz, M. Moncur, S. Gabriel, and a. N. a. Tosteson, "Assessing Values for Health: Numeracy Matters," *Med. Decis. Mak.*, vol. 21, no. 5, pp. 382–390, Oct. 2001.

[11] S. R. Schwartz, J. McDowell, and B. Yueh, "Numeracy and the shortcomings of utility assessment in head and neck cancer patients.," *Head Neck*, vol. 26, no. 5, pp. 401–7, May 2004.

[12] B. Fasolo and C. a. Bana e Costa, "Tailoring value elicitation to decision makers' numeracy and fluency: Expressing value judgments in numbers or words," *Omega*, vol. 44, pp. 83–90, Apr. 2014.

[13] A. Maynard and D. McDaid, "Evaluating health interventions: exploiting the potential," *Health Policy (New. York).*, vol. 63, pp. 215–226, 2003.

[14] E. Pereira, "Avaliação económica das tecnologias da saúde," *Rev. Ciências da Saúde da ESSCVP*, vol. 1, pp. 33–40, 2009.

[15] M. J. Drummond, Michael F. Sculpher, G. W. Torrance, B. J. O'Brien, and G. L. Stoddart, *Methods for the Economic Evaluation of Health Care Programmes*, Third Edit. Oxford University Press, 2005.

[16] D. Banta, "The development of health technology assessment," *Health Policy (New. York).*, vol. 63, pp. 121–132, 2003.

[17] "Health Technology Assessment international." [Online]. Available: http://www.htai.org/index.php?id=428.

[18] D. McDaid and R. Cookson, "Evaluating health care interventions in the European Union," *Health Policy (New. York).*, vol. 63, pp. 133–139, 2003.

[19] D. Banta, E. Jonsson, and P. Childs, "History of the international societies in health technology assessment: International Society for Technology Assessment in Health Care and Health Technology Assessment International.," *Int. J. Technol. Assess. Health Care*, vol. 25 Suppl 1, pp. 19–23, Jul. 2009.

[20] "INAHTA." [Online]. Available: http://www.inahta.org/.

[21] D. Hailey, "Development of the International Network of Agencies for Health Technology Assessment.," *Int. J. Technol. Assess. Health Care*, vol. 25 Suppl 1, pp. 24–7, Jul. 2009.

[22] "NICE." [Online]. Available: http://www.nice.org.uk/about.

[23] M. Drummond, "What are the HTA processes in the UK?," *Hayward Med. Commun.*, no. 4, pp. 1–8, 2009.

[24] "SiNATS." [Online]. Available: http://www.infarmed.pt/portal/page/portal/INFARMED/MEDICAMENTOS_USO_HUMANO/SINATS.

[25] J. Martins, J. Rodrigues, M. Antunes, F. Ferrador, I. Ramos, R. Ramos, C. Santos, S. Caldeira, J. Castro, M. Arriegas, P. Dias Almeida, H. Mota-Filipe, and E. Castro Alves, "Sistema Nacional de Avaliação de Tecnologias de Saúde para Portugal (SiNATS) - Criar o futuro," *INFARMED, I.P.*, pp. 1–88, 2014.

[26] D. McDaid, R. Cookson, A. Maynard, and F. Sassi, "Evaluating health interventions in the 21st century: old and new challenges," *Health Policy (New. York).*, vol. 63, pp. 117–120, 2003.

[27] A. Haycox, "What is health economics?," *Hayward Med. Commun.*, no. 4, pp. 1–8, 2009.

[28] C. R. Mitton, "Priority setting for decision makers: using health economics in practice.," *Eur. J. Heal. Econ.*, vol. 3, no. 4, pp. 240–243, Jan. 2002.

[29] S. Polinder, H. Toet, M. Panneman, and E. van Beeck, "Methodological Approaches For Cost-Effectiveness and Cost-Utility Analysis of Injury Prevention Measures," *World Heal. Organ. Reg. Off. Eur.*

[30] C. McCabe, "What is cost–utility analysis?," *Hayward Med. Commun.*, no. 2, pp. 1–6, 2009.

[31] L. N. Ferreira, "Avaliação Económica no Sector da Saúde," *dos Algarves*, no. 13, pp. 43–49, 2005.

[32] J. S. Hoch and C. S. Dewa, "An introduction to economic evaluation: what's in a name?," *Can. J. Psychiatry*, vol. 50, no. 3, pp. 159–66, Mar. 2005.

[33] B. Jönsson, "Changing health environment: the challenge to demonstrate cost-effectiveness of new compounds," *Pharmacoeconomics*, vol. 22, no. 4, pp. 5–10, Jan. 2004.

[34] R. J. Sperry, "Principles of Economic Analysis," *Anesthesiol. - Am. Soc. Anesthesiol.*, vol. 86, no. 5, pp. 1197–1205, 1997.

[35] L. Rudmik and M. Drummond, "Health economic evaluation: important principles and methodology," *Laryngoscope*, vol. 123, no. 6, pp. 1341–7, Jun. 2013.

[36] W. Schramm and T. D. Szucs, "State-of-the-art principles and practices of medical economics," *Haemophilia*, vol. 4, pp. 491–497, 1998.

[37] C. Phillips, "What is cost-effectiveness?," *Hayward Med. Commun.*, no. 2, pp. 1–8, 2009.

[38] L. de N. e Ferreira, "Utilidades, QALYs e Medição da Qualidade de Vida," *Assoc. Port. Econ. da Saúde*, 2002.

[39] P. J. Neumann, D. Greenberg, N. V Olchanski, P. W. Stone, and A. B. Rosen, "Growth and Quality of the Cost-Utility Literature, 1976-2001.," *Value Health*, vol. 8, no. 1, pp. 3–9, 2005.

[40] K. Stein, A. Fry, A. Round, R. Milne, and J. Brazier, "What Value Health? A Review of Health State Values Used in Early Technology Assessments for NICE," *Appl. Health Econ. Health Policy*, vol. 4, no. 4, pp. 219–228, 2005.

[41] M. C. Weinstein, G. Torrance, and A. McGuire, "QALYs: the basics.," *Value Health*, vol. 12 Suppl 1, pp. S5–9, Mar. 2009.

[42] C. Philips, "What is a QALY?," *Hayward Med. Commun.*, no. 4, pp. 1–6, 2009.

[43] D. L. Patrick, H. E. Starks, K. C. Cain, R. F. Uhlmann, and R. a. Pearlman, "Measuring Preferences for Health States Worse than Death," *Med. Decis. Mak.*, vol. 14, no. 1, pp. 9–18, Feb. 1994.

[44] J. Richardson, "Cost utility analysis: What should be measured?," *Soc. Sci. Med.*, vol. 39, no. 1, pp. 7–21, Jul. 1994.

[45] M. G. M. Weernink, S. I. M. Janus, J. A. van Til, D. W. Raish, J. G. van Manen, and M. J. Ijzerman, "A systematic review to identify the use of preference elecitation methods in healthcare decision making," 2014.

[46] G. H. Guyatt, D. H. Feeny, and D. L. Patrick, "Measuring Health-related Quality of Life," 1993.

[47] L. Fallowfield, "What is quality of life?," *Hayward Med. Commun.*, no. 5, pp. 1–8, 2009.

[48] G. W. Torrance, “Preferences for health outcomes and cost-utility analysis.,” *The American journal of managed care*, vol. 3 Suppl. pp. S8–20, May-1997.

[49] L. de N. Ferreira, P. L. Ferreira, and M. S. Gonçalves, “Ganhos em Saúde em doentes com cataratas,” *Rev. da Fac. Econ. da Univ. Coimbra*, vol. 23, pp. 35–52, 2006.

[50] P. L. Ferreira, “Criação da Versão Portuguesa do MOS SF-36 (Parte I),” *Acta Med. Port.*, vol. 13, pp. 55–66, 2000.

[51] I. Mcdowell, *Measuring Health : A Guide to Rating Scales*, Third. Oxford University Press.

[52] P. Dolan, C. Gudex, P. Kind, and a. Williams, “Valuing health states: A comparison of methods,” *J. Health Econ.*, vol. 15, no. 2, pp. 209–231, 1996.

[53] V. Shih, A. Chan, F. Xie, and Y. Ko, “Health State Utility Assessment for Breast Cancer,” *Value Heal. Reg. Issues*, vol. 1, no. 1, pp. 93–97, May 2012.

[54] C. Green, J. Brazier, and M. Deverill, “Valuing health-related quality of life. A review of health state valuation techniques.,” *Pharmacoeconomics*, vol. 17, no. 2, pp. 151–65, Feb. 2000.

[55] C. Gudex, P. Dolan, P. Kind, and a Williams, “Health state valuations from the general public using the visual analogue scale.,” *Qual. Life Res.*, vol. 5, no. 6, pp. 521–31, Dec. 1996.

[56] K. Tolley, “What are health utilities?,” *Hayward Med. Commun.*, no. 4, pp. 1–8, 2009.

[57] G. W. Torrance, W. Furlong, and D. Feeny, “Health utility estimation.,” *Expert Rev. Pharmacoecon. Outcomes Res.*, vol. 2, no. 2, pp. 99–108, Apr. 2002.

[58] M. Locadio, “A comparison of three valuation methods for temporary health states in patients treated with oral anticoagulants,” University of Amsterdam - Faculty of Medicine, 2004.

[59] D. R. Wright, E. Wittenberg, J. S. Swan, R. a. Miksad, and L. a. Prosser, “Methods for measuring temporary health states for cost-utility analyses,” *Pharmacoeconomics*, vol. 27, no. 9, pp. 713–723, 2009.

[60] J. Kopec, “A comparative review of four preference-weighted measures of health-related quality of life,” *J. Clin. Epidemiol.*, vol. 56, no. 4, pp. 317–325, Apr. 2003.

[61] G. W. Torralce, W. F. Ollg, D. Feellyl, and M. Boyle, “Multi-Attribute Preference Functions Health Utilities Index,” vol. 7, pp. 503–520.

[62] R. Brooks, “EuroQol: the current state of play.,” *Health Policy*, vol. 37, no. 1, pp. 53–72, Jul. 1996.

[63] J. Brazier, T. Usherwood, R. Harper, and K. Thomas, “Deriving a Preference-Based Single Index from the UK SF-36 Health Survey,” *J. Clin. Epidemiol.*, vol. 51, no. 11, pp. 1115–1128, Nov. 1998.

[64] S. Ali and S. Ronaldson, “Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods.,” *Br. Med. Bull.*, vol. 103, no. 1, pp. 21–44, Sep. 2012.

[65] R. T. Carson and J. J. Louviere, “A Common Nomenclature for Stated Preference Elicitation Approaches,” *Environ. Resour. Econ.*, vol. 49, no. 4, pp. 539–559, Jan. 2011.

[66] J. J. Louviere and E. Lancsar, “Choice experiments in health: the good, the bad, the ugly and toward a brighter future.,” *Health Econ. Policy. Law*, vol. 4, no. Pt 4, pp. 527–46, Oct. 2009.

[67] E. W. D. E. Bekker-grob, M. Ryan, and K. Gerard, “Discrete Choice Experiments in Health Economics: A Review of the Literature,” *Health Econ.*, vol. 21, pp. 145–172, 2012.

[68] J. a Whitty, E. Lancsar, K. Rixon, X. Golenko, and J. Ratcliffe, “A Systematic Review of Stated Preference Studies Reporting Public Preferences for Healthcare Priority Setting,” *Patient*, no. May, May 2014.

[69] P. Thokala and A. Duenas, “Multiple criteria decision analysis for health technology assessment.,” *Value Health*, vol. 15, no. 8, pp. 1172–81, Dec. 2012.

[70] V. Belton and T. J. Stewart, *Multiple Criteria Decision - An Integrated Approach*. Springer-Science+Business Media, B.V.

[71] G. B. Chapman, A. S. Elstein, T. M. Kuzel, R. B. Nadler, R. Sharifi, and C. L. Bennett, “A multi-attribute model of prostate cancer patients ’ preferences for health states,” *Qual. Life Res.*, vol. 8, pp. 171–180, 1999.

[72] C. a. Bana e Costa and J.-C. Vansnick, “A critical analysis of the eigenvalue method used to derive priorities in AHP,” *Eur. J. Oper. Res.*, vol. 187, no. 3, pp. 1422–1428, Jun. 2008.

[73] A. H. Harris, J. Youd, and R. Buchbinder, “A comparison of directly elicited and pre-scored preference-based measures of quality of life: the case of adhesive capsulitis.,” *Qual. Life Res.*, vol. 22, no. 10, pp. 2963–71, Dec. 2013.

[74] P. Krabbe, M.-L. Essink-Bot, and G. Bonsel, “The Comparability and Reliability of Five Health-State Valuation Methods,” *Soc. Sci. Med.*, vol. 45, no. 11, pp. 1641–1652, 1997.

[75] P. A. Ubel, G. Loewenstein, and C. Jepson, “Whose quality of life ? A commentary exploring discrepancies between health state evaluations of patients and the general public,” *Qual. Life Res.*, vol. 12, pp. 599–607, 2003.

[76] L. N. Ferreira, P. L. Ferreira, L. N. Pereira, and M. Oppe, “The valuation of the EQ-5D in Portugal.,” *Qual. Life Res.*, vol. 23, no. 2, pp. 413–23, Mar. 2014.

[77] L. N. Ferreira and P. L. Ferreira, “Utilização do SF-6D na medição das preferências dos portugueses : sistema de valores e normas da população dos 18 aos 64 anos,” *Rev. Port. saúde pública*, vol. 29, no. 2, pp. 108–115, 2011.

[78] P. L. Ferreira, L. N. Ferreira, and L. N. Pereira, “Contributos para a Validação da Versão Portuguesa do EQ-5D,” *Acta Med. Port.*, vol. 26, no. 6, pp. 664–676, 2013.

[79] L. N. Ferreira, P. L. Ferreira, L. N. Pereira, and M. Oppe, “EQ-5D Portuguese population norms.,” *Qual. Life Res.*, vol. 23, no. 2, pp. 425–30, Mar. 2014.

[80] P. L. Ferreira and L. de N. Ferreira, “A medição de preferências em saúde na população portuguesa,” vol. 24, pp. 5–14, 2006.

[81] L. N. Ferreira, P. L. Ferreira, L. N. Pereira, J. Brazier, and D. Rowen, “A Portuguese value set for the SF-6D,” *Value Health*, vol. 13, no. 5, pp. 624–30, Aug. 2010.

[82] L. N. Ferreira, P. L. Ferreira, and R. R. Baleiro, “Qualidade de Vida em Doentes com Artrite Reumatóide,” *Acta Reum. Port.*, vol. 33, pp. 331–342, 2008.

[83] L. N. Ferreira, P. L. Ferreira, and L. N. Pereira, “Comparing the Performance of the SF-6D and the EQ-5D in Different Patient Groups,” *Acta Med. Port.*, vol. 27, no. 2, pp. 236–246, 2014.

[84] L. N. Ferreira, U. Brito, and P. L. Ferreira, “Qualidade de vida em doentes com asma,” *Rev. Port. Pneumol.*, vol. XVI, no. 1, pp. 23–55, 2010.

[85] J. M. N. Catarino, “Qualidade de Vida Relacionada com o VIH,” 2010.

[86] M. Viswanathan and T. Childers, “Processing of Numerical and Verbal Product Information,” *J. Consum. Psychol.*, vol. 5, no. 4, pp. 359–385, 1996.

[87] L. M. Schwartz, S. Woloshin, W. C. Black, and H. G. Welch, “The Role of Numeracy in Understanding the Benefit of Screening Mammography,” vol. 127, no. 11, pp. 1023–1028, 2015.

[88] C. a. Bana E Costa, J.-M. De Corte, and J.-C. Vansnick, “Macbeth,” *Int. J. Inf. Technol. Decis. Mak.*, vol. 11, no. 02, pp. 359–387, Mar. 2012.

[89] C. Bana e Costa, J.-M. De Corte, and J. Vansnick, “M-MACBETH - Guia do utilizador,” pp. 1–56, 2005.

[90] C. A. Bana e Costa, L. Angulo-Meza, and M. Oliveira, “O Método MACBETH e Aplicação no Brasil,” *ENGEVISTA*, vol. 15, no. 1, pp. 3–27, 2013.

[91] C. A. Bana, J. De Corte, and J. Vansnick, “On the Mathematical Foundations of MACBETH,” *London Sch. Econ. Polit. Sci.*, 2004.

[92] N. Bansback, A. Tsuchiya, J. Brazier, and A. Anis, “Canadian Valuation of EQ-5D Health States: Preliminary Value Set and Considerations for Future Valuation Studies.,” *PLoS One*, vol. 7, no. 2, Jan. 2012.

[93] R. Norman, R. Viney, J. Brazier, L. Burgess, P. Cronin, M. King, J. Ratcliffe, and D. Street, “Valuing SF-6D Health States Using a Discrete Choice Experiment.,” *Med. Decis. Making*, vol. 34, no. 6, pp. 773–786, Sep. 2013.

[94] J. Richardson, A. Iezzi, and M. Khan, “The Self Time Trade-off (TTO) Instrument : Reliability and Survey Results,” *Monash Univ. Bus. Econ. - Cent. Heal. Econ.*, vol. 86, no. July, 2014.

[95] P. F. M. Krabbe, “A Generalized Measurement Model to Quantify Health: The Multi-Attribute Preference Response Model,” *PLoS One*, vol. 8, no. 11, Jan. 2013.

[96] P. Dolan, “Modelling valuations for EuroQol health states,” *Med. Care*, vol. 35, no. 11, pp. 1095–1108, 1997.

[97] CEISUC, “Avaliação de Ganhos em Saúde - EQ-5D (Versão Portuguesa).” 1998.

[98] J. Scott and E. C. Huskisson, “Vertical or horizontal visual analogue scales.,” *Ann. Rheum. Dis.*, vol. 38, no. 6, pp. 560–560, Dec. 1979.

[99] C. Gudex, “Time Trade-Off User Manual: Props and Self-Completion Methods.” The University of York - Centre for Health Economics.

[100] K. Burström, M. Johannesson, and F. Diderichsen, “A comparison of individual and social time trade-off values for health states in the general population.,” *Health Policy*, vol. 76, no. 3, pp. 359–70, May 2006.

[101] L. Lundberg, M. Johannesson, D. A. G. G. L. Isacson, and L. Borgquist, “Health-state utilities in a general population in relation to age , gender and socioeconomic factors,” *Eur. J. Public Health*, vol. 9, no. 3, pp. 211–217, 1995.

[102] C. Bardage, D. Isacson, L. Ring, and K. Bingefors, “A Swedish population-based study on the relationship between the SF-36 and health utilities to measure health in hypertension,” *Blood Press.*, vol. 12, no. 4, pp. 203–210, Jan. 2003.

[103] A. Fagerlin, B. J. Zikmund-Fisher, P. a Ubel, A. Jankovic, H. a Derry, and D. M. Smith, “Measuring numeracy without a math test: development of the Subjective Numeracy Scale.,” *Med. Decis. Making*, vol. 27, no. 5, pp. 672–80, 2007.

[104] I. M. Lipkus, G. Samsa, and B. K. Rimer, “General Performance on a Numeracy Scale among Highly Educated Samples,” *Med. Decis. Mak.*, vol. 21, no. 1, pp. 37–44, Feb. 2001.

[105] M. Voicu and A. Babonea, “Using the Snowball Method in Marketing Research on Hidden Populations,” *Challenges th Knowl. Soc. Econ.*, vol. 44, no. 2, pp. 102–104, 2007.

[106] M. Pocinho, “Estatistica - Volume I Teoria e exercícios passo-a-passo,” *ISMT*, vol. I, pp. 1–82, 2009.

[107] A. E. Attema, Y. Edelaar-Peeters, M. M. Versteegh, and E. a Stolk, “Time trade-off: one methodology, different methods.,” *Eur. J. Health Econ.*, vol. 14 Suppl 1, pp. S53–64, Jul. 2013.

[108] T. Arnesen and M. Trommald, “Are QALYs based on time trade-off comparable?--A systematic review of TTO methodologies.,” *Health Econ.*, vol. 14, no. 1, pp. 39–53, Jan. 2005.

ANEXO A

Mobilidade

1. Não tenho problemas em andar
2. Tenho alguns problemas em andar
3. Tenho de estar na cama

Cuidados Pessoais

1. Não tenho problemas em cuidar de mim
2. Tenho alguns problemas em lavar-me ou vestir-me
3. Sou incapaz de me lavar ou vestir sozinho(a)

Atividades Habituais

1. Não tenho problemas em desempenhar as minhas atividades habituais
2. Tenho alguns problemas em desempenhar as minhas atividades habituais
3. Sou incapaz de desempenhar as minhas atividades habituais

Dor/Mal-estar

1. Não tenho dores ou mal-estar
2. Tenho dores ou mal-estar moderados
3. Tenho dores ou mal-estar extremos

Ansiedade/Depressão

1. Não estou ansioso(a) ou deprimido(a)
2. Estou moderadamente ansioso(a) ou deprimido(a)
3. Estou extremamente ansioso(a) ou deprimido(a)

Figura 21. Sistema descritivo do EQ-5D (adaptado de [97].

ANEXO B

Tabela 16. Caraterização da amostra populacional que avaliou cada um dos quatro conjuntos de estados de saúde definidos.

			Grupo 1	Grupo 2	Grupo 3	Grupo 4
N			**53**	**52**	**53**	**52**
Variáveis sociodemográficas			**n (%)**	**n (%)**	**n (%)**	**n (%)**
Sexo	**Feminino**		35 (66.0)	32 (61.5)	34 (64.2)	33 (63.5)
	Masculino		18 (34.0)	20 (38.5)	19 (35.9)	19 (36.5)
Idade	**Média ± DP**		34.9 ±12.5	33.7 ± 12.7	34.9 ± 14.2	33.8 ± 13.3
	18-24		12 (22.6)	15 (28.9)	22 (41.5)	21 (40.4)
	25-30 *		14 (26.4)	12 (23.1)	3 (5.7)	8 (15.4)
	31-44		15 (28.3)	15 (28.9)	12 (22.6)	11 (21. 2)
	> 44		12 (22.6)	10 (19.2)	16 (30.2)	12 (23.1)
Nível de escolaridade	**Inferior ao Ensino Secundário**		1 (1.9)	1 (1.9)	1 (1.9)	0
	Ensino Secundário		10 (18.9)	12 (23.1)	17 (32.1)	13 (25.0)
	Superior ao Ensino Secundário		42 (79.2)	39 (75.0)	35 (66.0)	39 (75.0)
Estado civil	**Solteiro**		20 (37.7)	28 (53.9)	28 (52.8)	29 (55.8)
	Casado/União estável		33 (62.3)	23 (44.2)	23 (43.4)	22 (42.3)
	Divorciado/Separado		0	1 (1.9)	1 (1.9)	1 (1.9)
	Viúvo		0	0	1 (1.9)	0
Situação profissional	**Estudante**		15 (28.3)	20 (38.5)	22 (41.5)	19 (36.5)
	Empregado		32 (60.4)	29 (55.8)	24 (45.3)	23 (44.2)
	Desempregado *		1 (1.9)	0	0	6 (11.5)
	Reformado		2 (3.8)	2 (3.9)	2 (3.8)	3 (5.8)
	Doméstico		0	1 (1.9)	1 (1.9)	1 (1.9)
	Outra situação *		3 (5.7)	0	4 (7.6)	0
Agregado familiar	**1-2 elementos**		14 (26.4)	13 (25.0)	12 (22.6)	17(32.7)
	3-4 elementos		30 (56.6)	36 (69.2)	40 (75.5)	29 (55.8)
	5 ou mais elementos *		9 (17.0)	3 (5.8)	1 (1.9)	6 (11.5)
Relativamente ao rendimento mensal líquido	**Não limita a minha saúde**	Discordo	10 (18.9)	13 (25.0)	15 (28.3)	16 (30.8)
		Indiferente:	3 (5.7)	8 (15.4)	6 (11.3)	4 (7.7)
		Concordo	39 (73.6)	30 (57.7)	32 (60.4)	32 (61.5)
		NS/NR:	1 (1.9)	1 (1.9)	0	0
	Não limita a minha qualidade de vida	Discordo	13 (24.5)	17 (32.7)	19 (35.9)	18 (34.6)
		Indiferente:	2 (3.8)	4 (7.7)	4 (7.6)	4 (7.7)
		Concordo	37 (69.8)	30 (57.7)	30 (56.6)	30 (57.7)
		NS/NR:	1 (1.9)	1 (1.9)	0	0
	Não limita o acesso que preciso a cuidados de saúde	Discordo	10 (18.9)	18 (34.6)	20 (37.7)	16 (30.8)
		Indiferente:	3 (5.7)	3 (5.8)	2 (3.8)	4 (7.7)
		Concordo	39 (73.6)	30 (57.7)	31 (58.5)	32 (61.5)
		NS/NR:	1 (1.9)	1 (1.9)	0	0
	Não limita os cuidados médicos que a minha saúde necessita	Discordo	11 (20.8)	19 (36.5)	18 (34.0)	17 (32.7)
		Indiferente:	4 (7.6)	4 (7.7)	4 (7.6)	3 (5.8)
		Concordo	37 (69.8)	28 (53.9)	31 (58.5)	32 (61.5)
		NS/NR:	1 (1.9)	1 (1.9)	0	0
	Deixa-me confortável que conseguirei lidar com eventuais problemas de saúde que surjam	Discordo	15 (28.3)	19 (36.5)	20 (37.7)	23 (44.2)
		Indiferente:	3 (5.7)	5 (9.6)	6 (11.3)	4 (7.7)
		Concordo	34 (64.2)	27(51.9)	27 (50.9)	25 (48.1)
		NS/NR:	1 (1.9)	1 (1.9)	0	0
Doença crónica	**Sim**		8 (15.1)	8 (15.4)	13 (24.5)	10 (19.2)
	Não		43 (81.1)	44 (84.6)	38 (71.7)	41 (78.9)

	NS/NR		2 (3.8)	0	2 (3.8)	1 (1.9)
Relativamente à doença crónica, n (%)	**Não limitam a minha qualidade de vida**	Discordo Indiferente: Concordo	1 (12.5) 0 7 (87.5)	5 (62.5) 0 3 (37.5)	8 (61.5) 0 5 (38.5)	4 (40.0) 0 6 (60.0)
	Não me impedem de realizar as atividades habituais que as pessoas da minha idade realizam	Discordo Indiferente: Concordo	1 (12.5) 1 (12.5) 6 (75.0)	2 (25.0) 2 (25.0) 4 (50.0)	6 (46.2) 0 7 (53.8)	1 (10.0) 2 (20.0) 7 (70.0)
	Não aumentam o risco de eu vir a sofrer de outras doenças	Discordo Indiferente: Concordo	4 (50.0) 0 4 (50.0)	5 (62.5) 0 3 (37.5)	7 (53.8) 1 (7.7) 5 (38.5)	6 (60.0) 2 (20.0) 2 (20.0)
	Não diminuem os anos de vida que as pessoas da minha idade esperam viver	Discordo Indiferente: Concordo	3 (37.5) 0 5 (62.5)	3 (37.5) 1 (12.5) 4 (50.0)	3 (23.1) 4 (30.8) 5 (38.5)	3 (30.0) 2 (20.0) 5 (50.0)
Numeracia, n (%)	**0 respostas certas**		1 (1.9)	2 (3.8)	2 (3.8)	1 (1.9)
	1 resposta certa *		1 (1.9)	9 (17.3)	6 (11.3)	8 (15.4)
	2 respostas certas		15 (28.3)	9 (17.3)	13 (24.5)	18 (34.6)
	3 respostas certas		36 (60.4)	32 (61.5)	32 (60.4)	25 (48.1)

* Diferença estatisticamente significativa entre grupos a um nível de 5%