



# Measuring Patient Preferences

---

An exploratory study to determine how patient preferences data could be used in health technology assessment (HTA)

---

Project report

June 2019



# Measuring Patient Preferences

---

An exploratory study to determine how patient preferences data could be used in health technology assessment (HTA)

---

Project report

June 2019



# ACKNOWLEDGEMENTS

---

Myeloma UK and NICE acknowledge the input of all the researchers who have worked on this study to date: Dr Sarah Garner, Eric Low, Sarah Richards, Dr Luke Cowie, Dr Jayne Galinsky, Dr Rosie Lovett, and Dr Jacqueline Bouvy. We also thank the patients who guided and supported the study and those who took part in workshops across the UK.

# CONTENTS

---

About this report .....	6
Glossary .....	7
1. Executive summary .....	8
2. Introduction .....	9
3. Literature Review .....	10
3.1 Background .....	10
3.2 Patient-reported outcome measures, QALYs and patient preferences... ..	10
3.3 Stated preference methods .....	11
3.4 Use of patient preferences studies in different settings .....	13
3.5 Barriers to using patient preferences in regulatory & HTA processes ... ..	15
4. Patient Insights Research .....	17
4.1 Aims, objectives and study design .....	17
4.2 Survey results .....	17
4.3 Focus group results .....	21
4.4 Survey and focus group conclusion .....	23
5. Which preference elicitation method? .....	24
5.1 Rating, ranking, and trade-offs .....	25
5.2 The importance of design .....	26
5.3 Best-worst scaling .....	26
6. How might patient preference studies be used by NICE? .....	28
6.1 The potential value of patient preference data in HTA .....	28
6.2 Patient preferences and cost-effectiveness analysis .....	29
6.3 The potential problem of heterogeneity .....	30
6.4 Concerns related to the use of patient preferences studies in HTA .....	30
6.5 What role do HTA bodies have? .....	31
7. Conclusions .....	32
8. References .....	34
Appendix 1: Study Design and Methods .....	38
Appendix 2: Best-worst scaling (BWS) designs .....	40

# ABOUT THIS REPORT

---

This report comprises the findings of a two-year research project that was a collaboration between NICE's Science Policy and Research programme and Myeloma UK, who funded the project and who have demonstrated a strong commitment to this area of research.

The contents of this report do not reflect an official NICE position on the use of patient preferences in HTA. The nature of the work was exploratory and cannot be understood as replacing or complementing official NICE guidance or any of its methods guides.

## Authors:

Luke Cowie and Jacoline Bouvy, NICE

# GLOSSARY

---

<b>BWS</b>	Best worst scaling
<b>EMA</b>	European Medicines Agency
<b>EQ-5D</b>	EuroQol 5 dimensions (instrument to measure quality of life)
<b>DCE</b>	Discrete choice experiment
<b>FDA</b>	Food and Drug Administration
<b>IQWiG</b>	Institute for Quality and Efficiency in Healthcare
<b>HRQoL</b>	Health-related quality of life
<b>HST</b>	Highly specialised technologies
<b>NHS</b>	National health service
<b>NICE</b>	National Institute for Health and Care Excellence
<b>QALY</b>	Quality-adjusted life year
<b>SF-6D</b>	Short form 6 dimensions (instrument to measure quality of life)

# 1. EXECUTIVE SUMMARY

---

The aim of this project was to undertake research to explore how quantitative methodology for eliciting patient preferences might be used in HTA. Patient preference elicitation methods can be defined as methods for collecting and analysing data that allow quantitative assessments of the relative desirability of treatment attributes to patients. In short, these methods can provide insights into the relative weights that people assign to different treatment options available to them.

The project activities involved a review of the existing patient preferences literature, primary research with myeloma patients, and the engagement of a range of stakeholders in the context of a critical assessment workshop. Taken together, the findings from these research activities have provided an overview of the potential value of using quantitative patient preference data in HTA and identified a range of different applications to which these data could be usefully applied.

The patient viewpoint is essential for HTA because patients can best represent the experience of living with their conditions. HTA bodies such as NICE have well-established procedures for including the views of patients and other stakeholders into the recommendations they make. Information on patient preferences is most often anecdotal testimony obtained in advisory committee proceedings, through public meetings, or through consultation. While these contributions make sure the patient perspective is used in the decision-making process, it does mean that HTA consideration of patient testimony is usually qualitative. As a result, its influence on decision making is not always obvious to those who were not involved in the deliberation. There is growing interest in using methods such as quantitative patient preference data in HTA as they could improve the transparency of using patient input in HTA.

This project found that there is no 'one-size-fits-all' solution for generating patient preference data. Instead, the choice of method will depend on the specific research question, the size of the patient population to be studied, and resource constraints. It is clear, however, that the DCE stands out as the method that offers a robust approach for generating insights in the relative importance of different treatment attributes, and the trade-offs that patients are willing to make between attributes. But the DCE is not always likely to be an appropriate method, for various reasons. Therefore, other methods should be considered when speed, ease of comprehension or cost are significant factors. Furthermore, not every recommendation that an HTA body needs to make will necessarily benefit to the same degree from quantitative patient preference data. More work is needed to better identify what types of recommendations can most benefit from using these methods.



## 2. INTRODUCTION

---

The aim of this project was to undertake exploratory research to enable the development of quantitative methodology to incorporate patient preferences into HTA. Patient preference elicitation methods can be defined as methods for collecting and analysing data that allow quantitative assessments of the relative desirability of attributes to patients that differ among alternative options<sup>1</sup>. In short, these methods can provide insights into the relative weights that people assign to different treatment options available to them.

Methods for eliciting preferences have evolved out of different research disciplines, such as economics, consumer marketing and decision analysis, and have been used to understand a diverse range of choice behaviours and valuation questions, such as in the context of transport demand (Hensher 1994) and valuing environmental amenities (Adamowicz et al. 1994). Preference data are increasingly being accepted for guiding regulatory decision making. The EMA and IQWiG in Germany have identified choice experiments as a potential method for supporting licensing and reimbursement decision making (Mühlbacher et al. 2017, EMA 2012). The use of quantitative patient preference data is still relatively minimal in HTA, where patient input is usually captured qualitatively through opinion and consultation (Mott 2018). Capturing patient preferences in a more quantitative way could enable this information to be incorporated more formally into decisions alongside other evidence as part of the HTA decision making process.

The specific objectives of this research project were to:

1. Review the literature and current research activity relating to patient preferences, with a focus on the methodological literature in HTA and related disciplines as well as sectors other than health;
2. Gather insights from patient with myeloma (also known as multiple myeloma) and their carers about what aspects of patient experience and types of preferences should be captured;
3. Critically assess potential methods in terms of their applicability and suitability to the HTA context;
4. Prepare a proposal to obtain funding for developing the methodology further.

This project report will provide an overview of all research activities that have been completed in order to fulfil the project objectives. As some of the activities will be further reported through scientific publications, this report provides a condensed overview of the research conducted and discusses the findings and their implications for the applicability of quantitative patient preference elicitation methods to HTA.

---

<sup>1</sup> A framework for incorporating information on patient preferences regarding benefit and risk into regulatory assessments of new medical technology. Medical Device Innovation Consortium (MDIC) <http://mdic.org/pcbr/framework-report/>

## 3. LITERATURE REVIEW

---

### 3.1 Background

Medicines, medical technologies and healthcare systems all exist in order to bring health benefits to patients. Whilst clinicians have long recognised the importance of listening to patients, traditional models of health care organisation and delivery have historically exhibited a more paternalistic attitude towards patients. Critiques of medical paternalism have, in tandem with wider social, cultural and technological transformations, gradually led to the implementation of changes seeking to redress the perceived inequality of patients as stakeholders within contexts of healthcare research, delivery and decision-making (Anderson & McCleary 2015). Patient and public involvement is now common to a wide range of health-related spheres of activity, either within guidelines for “best practice” or explicitly required within health research or processes of healthcare technology regulation and evaluation. The issue of quantitative patient preferences fits into this wider movement towards greater patient participation and representation in decisions that directly impact on their care.

Health economists have long been familiar with the use of preference elicitation methods, particularly for the valuation of health states (Devlin et al. 2018). Researchers measure the preferences of patients or the public in order to create value sets that are an important component of the use of instruments such as the EQ-5D, that are frequently used in HTA where the EQ-5D can be used to measure and value HRQoL. Historically, methods such as the time trade-off or standard gamble have been used to elicit these preferences. But other methods, such as the DCE and BWS are increasingly being used (Clark et al. 2014). To date, other applications for quantitative patient preference elicitation methods, such as the generation of data for supporting healthcare decision making, remain relatively rare.

### 3.2 Patient-reported outcome measures, QALYs and patient preferences

In discussing patient preference elicitation methods, it is important to first clarify some closely related concepts and their terminology. Patient-reported outcome measures comprise disease-specific and/or general measures that collect data on symptom status, physical function, mental health, social function and wellbeing. Disease-specific measures are often more responsive to changes in health than measures that can be used across different conditions (generic measures), but have limited usefulness for cost-effectiveness analyses and other broader analyses because they may not effectively capture the effects of comorbidities and do not allow for comparisons across conditions (Feeny et al. 2013). Patient-reported outcome measures are intended to measure changes in health status from the perspective of the individual who is undergoing a healthcare, and so

they are increasingly used within clinical trials. Though there are similarities and areas of overlap, it is useful to distinguish between HRQoL measures and patient preference methods, which are a distinct class of methods that are increasingly being applied in medicine.

Instruments to measure and value HRQoL are fundamental to HTA in many countries and in countries that use cost-effectiveness analysis in HTA, generic HRQoL instruments tend to be preferred over disease-specific instruments. While generic instruments, such as the EQ-5D or SF-6D, are attractive for their generic nature that enables comparisons of different treatments across a range of conditions, they were not designed for preference evaluation. HRQoL might not capture all factors that are important to patients, providers or policy-makers, such as non-health outcomes or process characteristics. Such attributes may include things like waiting times and modes of access to healthcare, mode and frequency of treatment administration, and impacts on family members, for example. Furthermore, patient preferences studies can provide insights into *trade-offs* between important health outcomes, non-health outcomes and process attributes.

Type of measure	Health-related quality of life	Patient preferences
<b>Objective</b>	Measuring health status of the patient	Valuing health states, outcomes, or attributes of treatments
<b>Temporal focus</b>	Real time, before/after an intervention	Preferences are relatively enduring, though may change over time
<b>Domains</b>	Pre-defined domains (e.g. mobility; self-care; usual activities; pain/discomfort; and anxiety/depression)	Outcomes/attributes selected according to study objectives

Table 1: Differences between HRQoL and patient preferences

### 3.3 Stated preference methods

Patient preference data are typically collected via one of two approaches:

- **Revealed preference** studies, which rely on observations of actual choice behaviour in real-world settings, examining the choices and trade-offs that people have actually made;
- **Stated preference** studies in which respondents participate in hypothetical experiments to determine their priorities and preferences.

Stated preference methodologies are more commonly used in the context of healthcare research because health is not a ‘normal’ market with traditional pricing mechanisms. Revealed preference methods are only possible for existing services that are available in the market place where consumers make their own

decisions about what product to purchase. In contrast, the hypothetical nature of stated preference methods allows them to be used to value interventions that are either not currently available, or for treatment decisions that in practice will involve a physician who may or may not prescribe the intervention.

Most stated preference methods are based on random utility theory. Economists use a measure called an 'util' for the amount of satisfaction you receive from something. An util is an abstract concept because it isn't something that exists in the physical world like an inch or a kilogram, but rather, it represents one unit of satisfaction or happiness. Although we cannot measure utility directly, economists will assume that when people make a choice between different options, they will choose the option which will bring them most satisfaction (utility). This is called ordinal utility, and it assumes that when someone chooses A over B, the utility of A is greater than the utility of B. It is this assumption that underlines the majority of stated preference methods. It is important to note, in the context of treatment choice, that this is not necessarily the same as choosing the option that will provide the best *health* outcomes.

Three broad categories of ordinal stated preference methodology have been distinguished: *ranking*, *rating* and *choice-based* approaches (Ryan et al. 2001). Rating and ranking methods are attractive because they are relatively simple and inexpensive to implement and are familiar to respondents. However, the lack of explicit trade-offs between benefits and harms, or between different attributes of a treatment, make ranking and rating less suitable from a methodological perspective (van Til & IJzerman 2014). Choice-based approaches, such as conjoint analysis and DCEs have greater grounding in economic theory.

The DCE is the most common type of ordinal preference method used in health economics and health services research and is based on a well-tested theory of choice behaviour. In DCEs, individuals are expected to make trade-offs in the context of a series of imperfect alternatives with different attribute profiles. A DCE will assume that the alternative with the greatest utility will be chosen among a set of alternatives. The DCE is considered to be a more realistic representation of actual decision-making as it allows for the estimation of overall preferences for any given combination of attributes and is shown to be one of the most sensitive methods to elicit preferences.

Another related method that is being used more frequently is BWS. Confusingly, BWS actually refers to three different variants: attribute case, profile case and multi profile case (otherwise known as Types 1, 2 and 3 respectively). Each question in these BWS methods provides two data points: best and worst (or most preferred and least preferred). Of the three variants, multi profile case is the most complex in terms of the data that is provided, and is the one that most closely resembles the DCE. This BWS variant will be explored in more detail in section 5 of this report. Examples of attribute case and profile case are presented in Appendix 2.

Attribute	Option A	Option B
<b>Effectiveness</b>	50% reduction of symptoms	80% reduction of symptoms
<b>Treatment duration</b>	1 week	2 weeks
<b>Dosage regimen</b>	Daily pill	Pill twice a day
<b>Side effects</b>	Mild headache	Moderate headache

Table 2: An example of a fictional DCE choice set

Conjoint analysis is an umbrella term that incorporates several related stated preference methods. Conjoint analysis methods feature less commonly in the patient preference elicitation literature, in large part because these methods originated in mathematical psychology and were principally developed for use in market research to determine how people value different attributes that make up individual products or services. There is sometimes confusion regarding terminology, and researchers may refer to their work as conjoint analysis when actually they have conducted a DCE. Traditional conjoint analysis will typically use rating scales to indicate the desirability or attractiveness of the profile, or alternatively will ask respondents to rank profiles in preference order. Of particular interest for preference research with patients is a hybrid form called adaptive conjoint analysis that requires respondents to use a ‘self-explicated’ method to weight all possible attribute-levels in order to create a reduced list (hence adaptive) that can be included in the conjoint exercise. But it is interesting to note the growing body of evidence to suggest that the simple ‘self-explicated’ method, which is much quicker, easier and less expensive to design and analyse than conjoint, is equal to conjoint in terms of internal and predictive validity (Johnson 2017). This is potentially an important finding, particularly in terms of its usefulness for studies with larger numbers of attributes, and also for its reported improved reliability in smaller sample sizes (Sattler & Hensel-Borner 2000).

### 3.4 Use of patient preferences studies in different settings

Internationally, several organisations are proactively re-evaluating their methods and exploring the potential benefits of patient preference data. In the United States, the FDA has been active in responding to and shaping the new field of patient input science. As part of the 2012 Safety and Innovation Act, the FDA’s Centre for Drug Evaluation and Research initiated its Patient-Focused Drug Development initiative, that aims to obtain patients’ perspectives on disease conditions and available treatments using a more systematic and expansive approach. Also in 2012, the FDA Centre for Devices and Radiological Health issued guidance which stated that it would now consider “patient tolerance for risk and perspective on benefit” in its decision making. The guidance went on to describe how risk tolerance inevitably varies among patients and that this would impact individual

decisions regarding how much risk was acceptable for a specified benefit, and that the FDA would now, for the first time, consider quantitative evidence relating to this (Johnson & Zhou 2016).

The first study designed to obtain quantitative patient-preference evidence to inform regulatory approval decisions by the FDA Center for Devices and Radiological Health (CDRH) was conducted using a DCE with 540 obese adults. The survey was designed to estimate the minimum weight loss acceptable for a patient to receive a device with a given risk profile, and the maximum mortality risk tolerable in exchange for a given weight loss (Ho et al. 2015). The CDRH has taken a central role in encouraging and facilitating the use of patient preferences in regulatory decision making (Johnson et al. 2017), and has finalised guidance on how to incorporate patient preferences into future benefit-risk assessments (FDA 2016). This development has been widely interpreted as signalling a likely rise in the application of patient preference data to regulatory decision making in the near future (Levitan et al. 2017). In 2016 the FDA issued guidance on the potential use of patient preference data for evaluating the benefit-risk profile of devices, particularly where patient decisions can be considered to be 'preference sensitive' (FDA 2016). Patient decisions regarding treatment options may be considered to be preference sensitive when:

1. multiple treatment options exist and there is no option that is clearly superior for all patients;
2. when the evidence supporting one option over others is considerably uncertain or variable; and/or
3. patients' views about the most important benefits and acceptable risks of a technology vary considerably within a population, or differ from those of healthcare professionals.

Regulators in Europe are also moving towards improving their approach to involving patients in the regulatory process. Like the FDA, the EMA recognises that patient views of risk and benefit can differ from those of other stakeholders, and may vary between patients at different stages of disease, and so patients can offer important insights into the risk-benefit assessments that lie at the heart of their activities. Recent changes have been made to improve their patient involvement processes, and while the EMA recognise the potential benefit of eliciting and quantifying patient preferences for different treatment outcomes using innovative methods, this is not yet part of their practice (European Medicines Agency 2013). A European project funded through the Innovative Medicines Initiative called the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) investigated the usefulness of different methodologies for more effectively capturing and communicating patient risk benefit analyses for use in regulatory decision making. The use of DCEs for this purpose was among their recommendations (Hughes et al. 2016). The ongoing successor to this project is called Patient Preferences in Benefit-Risk Assessments during the Drug Life Cycle (PREFER) and is a five year public-private research project where academic researchers, HTA bodies, patient organisations, and the pharmaceutical industry work together to find out when and where patients want, can, and should

be involved in drug development. With a broader remit, including the field of HTA, this project will deliver recommendations on the use of patient preferences across the technology lifecycle by 2021 (de Bekker-Grob et al. 2017).

### 3.5 Barriers to using patient preferences in regulatory & HTA processes

The potential advantages of using patient preference data for supporting decisions about the risks and benefits of new treatments are tempered by legal, methodological and pragmatic barriers to their use in this context. Legal issues relate to the commissioning responsibility for collecting patient preferences data, the reliability of the data produced, and to the consequences of potential biases that may occur. The methodological limitations of the various methods used to elicit patient preferences must be explored and fully understood, including the potential for innumeracy, the use of cognitive shortcuts by respondents (heuristics), variation among subgroups, inert or inflexible preferences, measurement error and hypothetical bias (Egbrink & IJzerman 2014).

Potential barriers may also be more pragmatic and procedural in nature. Patient preference studies require a significant amount of expertise and experience to design and implement effectively, and can be time and cost intensive. The research design is crucial to ensure that the results of the elicitation method are to be relevant to the question being deliberated by decision makers. It is important that patient preference studies are seen to be impartial and independent. The methods selected for the preference study and the information it is seeking to elicit should be pre-agreed between the company, the regulator and by the relevant patient group(s). It should be designed to fit within existing processes of decision-making such that it does not cause delays or disruptions, and should be presented in an appropriate form at the most relevant stage of the decision making process (van Til & IJzerman 2014). Existing scientific advice procedures that regulatory agencies and HTA bodies already offer to manufacturers for receiving advice on their clinical development plans could also be used for advice on study protocols for patient preferences studies, especially if manufacturers are planning to use the results of these studies as part of their submissions to regulatory agencies and/or HTA bodies.

The barriers identified above are likely to apply equally within a regulatory framework as in the context of HTA decision making. But there is also an important difference in emphasis: whereas regulation is unambiguously oriented towards the analysis of risk-benefit from the perspective of patients, HTA bodies must consider these aspects as part of a broader framework. The remit of HTA bodies across Europe differs, but many will have aims related to securing the best “value for money” within an assumed context of scarce resources. As a result, a wider decision-making perspective must be incorporated. This has meant in some jurisdictions that concerns about “distributive justice” (Rawlins 2009) and the preferences of the taxpaying public take priority over the preferences of individual patients or groups of patients within recommendations informed by HTA in

healthcare systems with collective provision. This does not mean that HTA bodies are not interested in patient preferences, but that their recommendations need to take additional considerations into account.

The different kinds of value that must be taken into account within HTA was discussed at a Health Technology Assessment International (HTAi) Policy Forum (Henshall & Schuller 2013):

- **The patient perspective:** Predominantly focussed on length and quality of life. Patients' views on value tend to reflect the whole of the care pathway and the way that they experience treatments within the context of the wider healthcare system, rather than of the treatment in isolation.
- **The societal perspective:** The need to maximise the "public good". But the views of the public are inherently complex and challenging to define. They will change over time and in response to media coverage of specific health issues, for example.
- **The health system perspective:** The allocation of resources to treatments of proven value, and assessing the value gained for a particular patient group against the opportunity cost of that provision to other patients.
- **The industry perspective:** Closely aligned with value for the patient because profits are ultimately derived from matching innovation to patient needs and preferences.

The wider aspects of value for patients, caregivers and society are not always measured or captured in a quantitative way in decision-making processes informed by HTA.

The potential fit of quantitative patient preferences data within HTA is therefore not without challenges. In risk-benefit decision making by regulatory agencies such as the FDA and the EMA, the interests of patients are primary, but in recommendations made by HTA bodies the needs of a specific group of patients who could benefit from a new technology must be balanced against the needs of the (typically unknown) other patients who will bear the opportunity cost, and HTA bodies also need to take non-patient factors into account in their recommendations. Ensuring a level playing field for using patient preferences data in all recommendations made by HTA bodies might prove problematic as the quantitative nature of patient preferences methods mean that it will never be possible to use these studies for all patient groups. A sufficient sample size, for example, is inevitably needed to ensure meaningful analyses of the data can be performed, but this means that HTA for technologies indicated for (very) rare conditions might never be able to incorporate quantitative patient preference data such as those derived using DCEs due to insufficient patient numbers being available. Ultimately, this means that the role of patient preferences data in HTA might fall more naturally as a form of evidence that can usefully sit alongside other types of evidence (such as economic modelling), rather than patient preferences being incorporated into economic models directly.



# 4. PATIENT INSIGHTS RESEARCH

---

## 4.1 Aims, objectives and study design

We performed a study to better understand the aspects of the patient experience and types of preferences that are important to patients but are not necessarily captured well in current HTA processes. Specifically, we aimed to answer the following question:

- From the perspective of people with myeloma, what are the patient preferences that should be considered by organisations such as NICE when conducting appraisals and/or producing guidance on treatments for the condition?

Furthermore, we also wanted to study whether specific patient characteristics or experiences were associated with particular preferences, and if so, what the implications would be for eliciting patient preferences for HTA. In addition, we wanted to explore whether the identified attributes could help us to identify the most appropriate preference elicitation method to capture them. Finally, we studied the practical considerations for conducting a patient preference study with myeloma patients.

We used two complementary research methods: a survey of myeloma patients and a focus group comprising equal numbers of myeloma patients and carers. A more detailed methods description can be found in Appendix 2.

## 4.2 Survey results

### 4.2.1 Survey demographics

Our sample consisted of 97 respondents in total: 55% were male and 45% were female. 72% of respondents were between 56 and 75 years old. These respondents were primarily living at home, with 77% living with at least one member of family, and 13% living alone. We found that a majority of respondents felt that they did not require any care, but 40% received some form of care, mostly from family members. Approximately 78% of respondents were either retired or currently unable to work due to ill health, and only 16% were currently employed.

Almost half of the respondents were actively receiving treatment at the time of the survey. Only three respondents had yet to receive any treatment for their myeloma. Nearly 32% had experienced a single course of treatment, and 38% had experienced three or more courses of treatment. Approximately 21% of respondents had started their first course of treatment one year previously, and 41% had started their first course of treatment within the past three years. 39% of respondents had started their treatment more than three years previously.

Nearly 60% of respondents reported having at least one other health condition, with nearly 20% having one additional condition, and 17% having two or more conditions in addition to myeloma<sup>2</sup>.

#### 4.2.2 Survey themes

The survey contained several open-ended questions that offered respondents the opportunity to input free text in order to report their experiences of living with myeloma, and to report the outcomes and other aspects of treatment that were important to them. Responses varied markedly in length, style, and content. These responses were coded using the methodology described in Appendix 2 and the frequency of each code is shown in the following figures. These coding frequencies can be interpreted as an indicator of the relevance and/or importance to this patient sample.

The survey asked respondents about the impact of myeloma on their lives, and on their lives and on those of their family and/or friends:

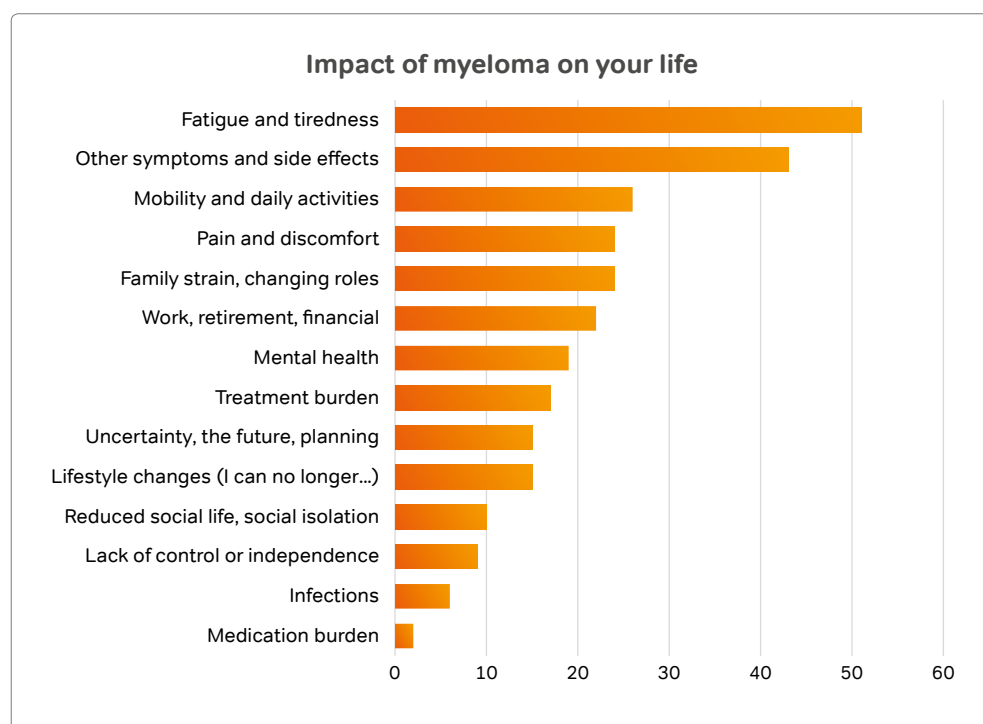


Figure 1: Impacts of myeloma on the lives of respondents

<sup>2</sup> By 'course of treatment' we specified this to mean the period of time spent taking the same chemotherapy regimen from beginning to end. If the respondent paused treatment for a short time and then started again, that would count as just one course. If the dose was changed, that would count as just one course. But if they stopped one regimen and started a new regimen with different drugs, that would count as two courses.

The survey also asked respondents “What are the most important good effects (or characteristics) that you would want from any treatment for your myeloma?”:

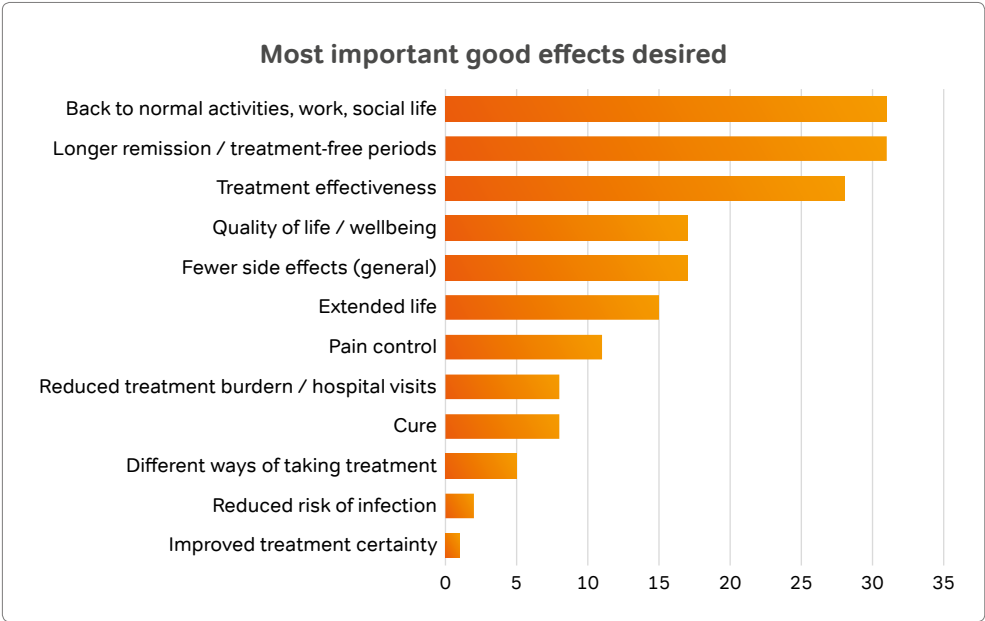


Figure 2: Treatment effects most desired by respondents

The survey also asked “What are the most important bad effects (or characteristics) that you would want to avoid from any treatment for your myeloma?:

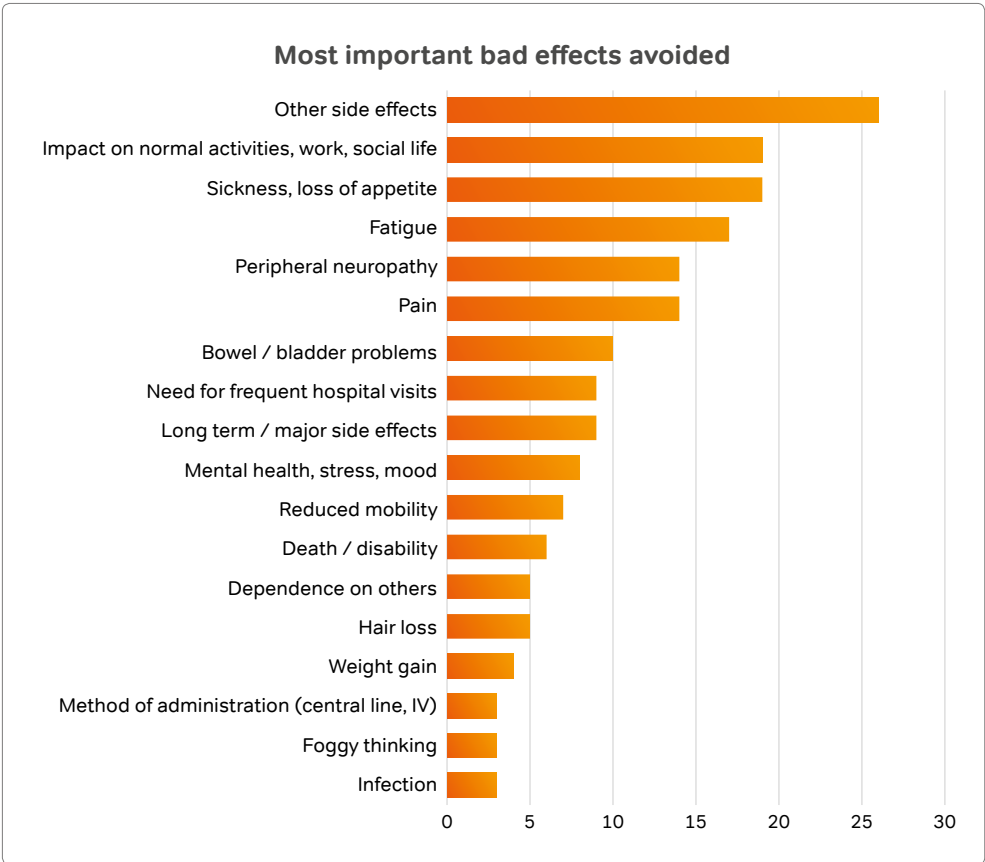


Figure 3: Treatment effects least desired by respondents

### 4.2.3 Survey discussion

The most commonly reported impact that myeloma had on the respondents' lives (*Figure 1*) was 'fatigue and tiredness', followed by side effects as a general category of experience. We identified a number of closely related themes that describe areas of deviation from "normal life" as a result of living with myeloma. This category included themes such as 'lack of control or independence', 'family strain, changing roles', 'lifestyle changes', 'work, retirement, financial', 'uncertainty, the future, planning', and 'reduced social life, social isolation'. These themes identify the clear impact of myeloma and its treatment on a wide range of issues that might impact quality of life and wellbeing. These themes clearly illustrate experiences that are important to patients and their families, some of which might not be effectively captured in a HRQoL instrument. Fatigue, for example, is a highly relevant symptom for myeloma patients. A generic HRQoL will measure changes in fatigue indirectly through its impacts on health dimensions included in the instrument. For example, if patients experience fatigue, this would impact their ability to carry out daily activities or their ability for self-care. However, HRQoL instruments were not specifically designed to measure preferences or trade-offs between different treatment options, and therefore, other types of evidence – such as patient preferences data – could provide additional insights into the relative preferences of patients for treatments that would improve fatigue.

We asked respondents about the most important positive effects that they would desire from a treatment for myeloma (*Figure 2*). One of the most commonly coded themes was 'treatment effectiveness', and this demonstrates, unsurprisingly, that respondents desired a treatment that will control their disease. Another commonly mentioned theme was 'longer remission/ treatment-free periods'. The reality of living with myeloma is that there is a finite number of treatment options available, and each of these has time-limited effectiveness for individual patients. Without new treatments becoming available, patients are limited in how long they can be effectively treated for myeloma. In related themes, respondents mentioned desiring 'extended life', 'fewer side effects', and improved 'quality of life/ wellbeing'. These factors are clinically driven and therefore are likely to be captured in estimates of clinical effectiveness that are part of HTA.

Respondents were also asked about the most important negative effects that they would be keen to avoid from a treatment for myeloma (*Figure 3*). 'Long term/ "major" side effects' was coded as a general theme to categorise responses where adverse reactions were described as long-lasting or permanent, or significantly damaging to the body, suggesting that the severity or consequences of side effects mattered more to patients than what exact side effect they would experience. In designing a DCE for myeloma patients this could be reason to include side effect severity as an attribute with attribute levels describing side effect severity, rather than including specific side effects as separate attributes. Explicit side effects that were commonly mentioned include 'sickness/ loss of appetite', 'bowel/ bladder problems', 'peripheral neuropathy', but also 'weight gain', 'hair loss', 'foggy thinking' and 'infection'. The breadth of outcomes discussed by respondents is a very useful indication of those areas that have important ramifications to the way

in which respondents must adapt to their condition and associated treatments, the impact that these have on their lives, and what aspects of treatment options would be important to capture in a DCE for patient with myeloma.

## 4.3 Focus group results

The focus group was an opportunity to facilitate discussion of the same questions asked in the survey, and thus to enable patients and carers to communicate their views and to compare and contrast their experiences with one another. It allowed us to reflect on the results of the survey and to ask focus group respondents to offer some of their own analysis of these results by discussing whether these data “made sense”, and whether or not there was anything they felt warranted particular mention, for example, if there was anything that surprised them about the results. We used the same process of thematic coding for the focus group as was used previously for the survey, and the final selection of thematic categories was very similar in terms of the breadth of issues that were discussed among the participants. The following section will give an overview of the thematic categories that were most prominent in the analysis, and/or most pertinent to the research questions.

### Dependence/independence

With three sets of partners in the focus group, the issue of relational dynamics between patient and carer were discussed at various points throughout the session. For the most part respondents seemed to agree that independence was threatened by myeloma, for various reasons. Respondents discussed that the lack of independence has knock-on impacts in terms of mental health and concerns for the future, in terms of family finances and coping after bereavement. This discussion served to underline the strength of impact that this disease has on the patient’s ability to self-care, and by extension, on their sense of self as someone who can function independently without support. The effects of these changes on important relationships was elaborated elsewhere in the focus group.

### Work, retirement, financial

Several respondents described how they were no longer able to work due to their illness and consequently were dependent upon other sources of income or financial support. One carer respondent described how they also had taken early retirement in order to be able to more effectively care for their spouse. The focus group also discussed the uncertain nature of the condition and how it makes it difficult to predict what the needs of the patient will be at various points in the future. In addition, patients could have prolonged periods of remission in which it may be possible to return to work.

## Impact on carers

The task of caring for a patient with myeloma can cause carers to experience difficulties of their own, particularly when there are no other family members in the household to contribute to caring or other necessary activities. Several respondents echoed the negative repercussions for the lives of those who live with and care for them. The issue of holidays was one that emerged quite strongly during the focus group as something that negatively impacted equally upon partner carers and patients. Holidays are usually one of the few opportunities that people have to spend extended periods of quality time together, but clearly this opportunity is not so straightforward with a condition such as myeloma.

## Medication/treatment burden & alternative modes of treatment

Focus group respondents did not go into much detail about the burden of treatment associated with intensive or prolonged periods of hospital admission, as was strongly suggested by respondents to the survey. What was discussed, however, were the closely related themes of medication burden and alternative modes of treatment. Respondents referred to the high number of tablets that they were taking (or had previously taken) on a daily basis. For some this was between 30 and 50 per day, and for one respondent in particular this was made even more challenging due to being on a liquid restricted diet at that time. Alternative ways of taking medication were discussed in the context of the potential for significant improvements to people's lives: one respondent mentioned that they forced themselves to overcome an extreme dislike of injections in order to be able to self-medicate at home rather than have to make regular trips to hospital to receive blood transfusions.

## Most important treatment outcomes

There was general agreement among respondents that longer remission periods and treatment effectiveness more generally were the most important outcomes. However, the focus group discussions made it clear there were differences in the relative desirability of specific outcomes with patients in the group, as different respondents had different perspectives on what the most important outcomes were for them personally. We found that differences in how the condition had been experienced by the patient would often be the reason for differences in preferred outcomes, for example the extent to which the respondent has experienced pain or other side effects. But even here, it was possible to see that personal attitudes will have an effect on decisions about pain control. Within the group it became clear that some were prepared to accept greater risk in order to benefit from improved pain management.

## Preference elicitation method suitability

The focus group discussed some of the relative merits of different patient preference elicitation methods. Four of these were explained to the group in simple terms using printed examples to illustrate the format of the question and to outline some of the characteristics that allow us to differentiate between them

with regards to aspects such as their relative cognitive burden, length of time to complete, and the kind of data that they produce. These methods were ranking, point allocation, DCE and BWS.

It was generally agreed that ranking and point allocation offered the more simple approaches to collecting preference data. Some respondents clearly indicated a preference for point allocation because strength of preference can be expressed. Others expressed a concern that point allocation might be more difficult to complete, because of the difficulty that some people have with numeracy. Another respondent made the point that some of the attributes listed in the provided example did not apply to them as they had not experienced them themselves, and another echoed the sentiment that because of the variability in how myeloma is experienced, some patients may not know that particular symptoms or side effects are associated with the disease or available treatments, if they have not experienced them personally. These reactions illustrate that in conditions where patient experiences are quite diverse, it might be more difficult to design a study that has equal validity across a patient population.

Respondents felt that non-engagement might be a problem if the measure takes more than 10 minutes to complete. In addition to the issue of non-engagement, there was also that of the cognitive burden associated with a particular method. In reaction to the two example methods presented that used a multiple question format using different attributes and attribute levels (DCE & BWS), some in the group suggested that prolonged concentration and the ability to continue to discriminate accurately between options might become an issue for some people.

## 4.4 Survey and focus group conclusion

The results of the survey and focus group of patients living with myeloma and their carers provided insights into how myeloma impacts the lives of patients and their carers. In addition, it demonstrates the need for qualitative research to inform the development of a quantitative patient preference study, such as a DCE, to ensure study designs are aligned with the experiences, needs, and preferences of the population to be studied. The focus group discussions brought about several aspects of treatments that matter to patients and their carers, such as treatment burden, treatment delivery, and non-health impacts, where a patient preference study could provide insights into the relative preferences of patients, especially if there would be substantial differences between treatment options available to patients. Furthermore, the focus group demonstrated that experiences among patients could differ substantially. This illustrates the possible limitation of relying solely on patient testimony in HTA processes for conditions where substantial heterogeneity among patients exists and where quantitative methods could provide better insights into the impacts of different treatment options on patients.

## 5. WHICH PREFERENCE ELICITATION METHOD?

---

No preference elicitation method is perfect, and therefore the issue of which method to select has been characterised as being a trade-off between cognitive burden for participants on the one hand and the possibilities for analysis on the other (Schmidt et al. 2016). Simple methods are easier to understand for participants but offer fewer options for analysis. Preference elicitation methods differ in relation to their modes of administration, underlying assumptions, scale of measurement, time allowed for reflection, use of props and supporting material, the extent to which they capture trade-offs and ease of comprehension (Ali & Ronaldson 2012). Other considerations include the present level of knowledge about patient preferences from other sources, how well the population of interest can be defined and accessed, the budget, expertise and time available to conduct the study, and the expected or available sample size (Ho et al. 2016, Tervonen et al. 2017). As a result, the question of the most appropriate method to use in HTA is not straightforward.

Choice-based methods such as the DCE have long been favoured over ranking and rating approaches (Ryan et al. 2001) and are commonly used in healthcare settings (Ali & Ronaldson 2012). But it is not appropriate to compare different methods head-to-head as though they are equivalent: the differences in presentation, framing, and overall methodology mean that they are fundamentally different and can be used for different objectives. Rating exercises, for example, are more appropriate for eliciting attitudes towards individual attributes, and DCEs are suited to understanding the trade-offs involved when making a decision, or when prioritising multiple aspects of health or healthcare (Wijnen et al. 2015).

Within the context of HTA, the selection of the most appropriate method for quantifying patient preferences depends on a range of contextual factors in addition to methodological considerations as described above. First, there needs to be substantial added value to the use of a patient preferences study. This is more likely to be the case when there are multiple treatment options available that have distinct characteristics. For example, if the HTA is of a technology indicated for a patient population where no treatment options exist, where the technology under evaluation is clearly very effective, or when there are several nearly identical treatment options available to patients, it is much less likely that the results of a quantitative patient preferences will impact the recommendation. However, if the technology under evaluation is very different from its comparator, for example when comparing pharmaceutical treatment to an operation, it might be very useful to decision makers to have the results of a well-designed and conducted patient preferences study that gives insights in how patients or groups of patients would value such treatments and their different attributes.



## 5.1 Rating, ranking, and trade-offs

Rating and ranking exercises are often regarded to be the simplest approaches to eliciting patient preferences (or priorities). Using the rating method (sometimes referred to as simple direct weighting), researchers will ask respondents to rate each outcome, or attribute of treatment, on an ordered Likert-type scale. Respondents will then indicate whether each item is 1=not at all important, 2=somewhat not important, 3=neutral, 4=somewhat important or 5=very important. But because respondents are not forced to make trade-offs, there is no motivation to think about the relative importance and this results in a tendency towards rating everything as important. A more fundamental problem is that ratings assume utility to be a cardinal construct. This assumption implies that a particular score measures the same utility across individuals and a given difference in score represents the same difference in utility across all situations. In terms of predicting choice, an ordinal construct that implies “better” or “worse” is more appropriate (Mühlbacher et al. 2016a).

Ranking is methodologically superior to rating because it is in accord with microeconomic theory of the ordinal nature of utility. Ranking exercises require respondents to rank a list of outcomes or important attributes in order from the most important to the least important. Whereas rating allows attributes to be equal in importance, ranking (usually) allows only one item to be allocated to each level, and thus forces respondents to prioritise the attributes. A weakness of this approach is that by forcing this choice process, this method does not allow for things to be valued equally. The simple hierarchical ordering also does not tell you anything about the relative importance of attributes.

Other methods can provide information on relative importance, whilst remaining relatively uncomplicated. One example is point allocation, in which respondents are asked to distribute a fixed number of units (often 100) between the listed set of outcomes or attributes according to perceived importance. Such an approach allows the respondent to allocate equally across all items if they should wish to, but also to rank them by allocating the units according to the strength of their preferences. This method can be adapted to context by using different scenarios to guide the respondents to allocate the units in a way that is appropriate to the research question being asked.

Trade-offs, whether generated by DCEs or other methods, are generally accepted as the most useful kind of patient preference information, because they help to determine the extent to which different attributes contribute to the total utility expected to be experienced by individual patients, or patient subgroups. BWS methods are comparatively recent addition to the stated preferences toolkit, and are increasingly being used to elicit preferences in healthcare (Long Cheung et al. 2016), and the relative value of these compared to the more traditional DCE should be considered.

## 5.2 The importance of design

There is often a trade-off between simplifying the decision problem and losing a degree of realism in designing a patient preference study. The design of any study type is a crucial factor in its relative acceptability and cognitive burden. The selection of appropriate attributes and levels should include consideration of what is feasible for respondents. In general, the more attributes that are included, the more challenging it is for respondents to consider them accurately. With DCEs in particular, a higher number of attributes and levels requires a greater number of choice questions, and this increases the risk of respondent fatigue and reduced data quality. Furthermore, respondents may not interpret the attributes and attribute levels in exactly the way that was intended by the researchers: depending on the context and their own experiences, respondents may interpret qualitative labels such as 'mild', 'moderate' and 'severe' differently (Mühlbacher & Johnson 2016).

DCEs force respondents to make a choice between two options (unless an 'opt-out' is included in the design) and therefore must be carefully designed using attributes that have been identified as being important to patients. Otherwise, DCEs can elicit counter-intuitive preferences that patients might not recognise as being their own. One example is the apparent willingness of some patients to trade some efficacy in IVF treatment for some improvement in the attitudes of staff. It is findings such as these that give credence to the idea that patient preferences are not so much revealed by these methods as generated by their specific design (Wainwright 2003). The inclusion of an 'opt-out' choice could prevent this from happening, but introduces more complexity in the analysis of the data and might have consequences for the sample size needed to generate meaningful results.

Checks for internal validity and consistency can be incorporated into DCEs, such as the dominance test, in which one of the alternatives in a choice scenario is clearly more desirable than the other. Another is the test-retest approach, whereby the same question is repeated later to check for consistency between responses. Although DCEs are often preferred because of a presumed similarity to the way in which people make decisions in the real world, the fact that some respondents 'fail' these tests and are often referred to as 'irrational' respondents indicates that the cognitive burden may be too high for some respondents. It has been argued that evidence of this kind of 'irrational' behaviour may suggest that choice behaviour in DCEs may not be consistent with utility theory (Ali & Ronaldson 2012). Notwithstanding, it demonstrates the importance of design considerations for patient preferences studies and the need to include checks that allow testing of the data validity.

## 5.3 Best-worst scaling

BWS methods are a comparatively recent development in the field of stated preference research, and though there are few evaluations of BWS methods in the literature, their use is steadily growing and there are several reasons why multi-profile BWS might be considered over the use of a DCE. Proponents argue that

it potentially obtains more information from respondents whilst exposing them to a lower cognitive burden, which may in turn have a positive influence on data quality and response rate (Whitty et al. 2014). The BWS methods all share the characteristic of asking respondents to make two choices per scenario, and so each time the respondent selects two items with maximum differences in utility. For this reason multi-profile BWS tends to have greater statistical efficiency than DCEs, and fewer question sets may be required.

A criticism of the BWS methods is that they are often not very realistic in terms of the way that people think about alternative options and make decisions. When buying a car, for example, you will not usually look at all the many varied options and select the 'worst' car, or the 'worst' attribute-level of a particular car. DCEs are arguably more natural, in that respondents are simply asked to make a positive choice between two options, which is something that we are all used to doing in everyday life. But there are situations in which choice is constrained, perhaps due to there being a very limited number of options to choose from. In such cases, people might more naturally consider what the 'least best' (i.e. 'worst') option is and exclude this in order to simplify the choice. Similarly, in some conditions (such as terminal cancer, for example) almost every option is a poor one. In these cases, the respondent is possibly more likely to be looking for the worst option, and so BWS might be more appropriate in some conditions more than others (Mühlbacher et al. 2016b).

## 6. HOW MIGHT PATIENT PREFERENCE STUDIES BE USED BY NICE?

---

In December 2017 we organised a critical appraisal workshop to discuss the application of patient preference methods within NICE processes. The workshop brought together a wide range of stakeholders with differing perspectives on the potential use of quantitative patient preference data in HTA, including NICE programme staff and representatives from industry, patient advocacy groups, and academia. This chapter reports the main themes that were raised in those discussions.

DCEs are the most well-tested and reliable approach for eliciting patient preferences that includes trade-off information. But these are very costly, time consuming and may pose a relatively high cognitive burden to some patients. BWS methods may offer some opportunity to reduce this cognitive burden by attaining a higher level of efficiency when eliciting preferences with the same (or fewer) number of choice sets. Where there is not a strong need for complex trade-off data, where small sample sizes are expected, where patient populations might be expected to struggle with complex choice-based approaches, or where there are acute time or other resource constraints, other approaches such as point allocation or a simple 'self explicated' method might be appropriate alternatives for eliciting patient preference data. The workshop focused on discussing the usability of both methods within the NICE context.

### 6.1 The potential value of patient preference data in HTA

There was a general consensus that patient preference elicitation methods are potentially beneficial in the context of HTA. The point was made that NICE depends upon high quality data from clinical trials in order to be able to make its recommendations, but that these data are not always optimal for informing all recommendations. Patient preferences data could potentially provide additional support that would enable decision makers to make more informed recommendations in cases where the clinical trial data in isolation might not provide a clear demonstration on precisely what the value proposition is for patients.

NICE is having to evaluate drugs earlier than in the past, with marketing authorisations being granted by regulators earlier resulting in the clinical data to support them being more limited. One attendee suggested that this further supports the use of patient preference data in NICE processes to make the evidence base more complete than it would otherwise be at that stage. The example was given of the HST programme at NICE, who already make efforts

to draw upon other sources of data because the trial data does not necessarily provide clear evidence around what patients want and need from new treatments. This view was supported by another attendee who commented on this from a patient advocacy perspective. Speedier access to new treatments is partly being driven by patient advocacy organisations, and these organisations also recognise the need for additional research to support the phase 1 and 2 data, because without quality data at that stage there is the risk that the value proposition will not be adequately demonstrated.

## 6.2 Patient preferences and cost-effectiveness analysis

Within the workshop discussions there was recognition of some potential barriers and areas of uncertainty that would need to be addressed if stakeholders are to be convinced that the potential value of patient preferences data can be realised in practice. The point was made that methods for patient preferences studies are not new, and so it is first important to recognise the existence of any cultural and organisational barriers that may have contributed to preventing the translation of extensive existing academic work in this field into HTA policy until now. One attendee went on to suggest that HTA bodies should be wary of a legacy commitment to a single methodology, arguing that it is important to look beyond NICE's current methods and make sure that we are measuring what is really important to patients.

There was some discussion about whether or not patient preference data should be used to directly value the health states generated by HRQoL instruments, or otherwise used in generating or valuing QALYs in cost effectiveness analysis. One attendee commented that there are limitations to the methods currently preferred by NICE for calculating QALYs, and that there might be ways in which patient preferences could be incorporated directly in cost effectiveness analysis. But this would constitute a substantial change to the methods that NICE currently prefers, and the use of preference elicitation methods to directly value health states would be in contradiction with the current requirement that these valuations should reflect the preferences of the general population.

The issue of equity in the context of using patient preferences data directly in calculating QALYs was also raised by several attendees: if patient preference data are only used for some appraisals and not others, then the patients using other treatments and services within the NHS that are at risk of disinvestment also have preferences that should perhaps be valued higher too. It was also suggested that in the near future it is likely that some disease areas would generate a large quantity of patient preference data, and other disease areas would have very little, if at all. It is also therefore a concern that the playing field might not be level, and that apples should not be compared to oranges when making comparative decisions.

Some attendees considered that the inclusion of patient preferences directly within the cost effectiveness analysis was likely to be problematic and a step too far. These attendees advocated the idea that patient preference data would be most valuable as an additional source of evidence alongside the current evidence NICE considers as part of its processes, and would be particularly valuable in those cases where it is recognised that patient preference data could provide insights that would not be adequately captured within NICE's current methods, for example, when two distinctly different treatment options need to be compared.

## 6.3 The potential problem of heterogeneity

There was some debate about how the heterogeneity of preferences among patients with the same condition might impact the use of patient preferences studies. One attendee argued that when considered alongside trial data on the various risk estimates and specific side effects of the technology under consideration, quantitative patient preferences potentially offer valuable insights how the value proposition for the technology might differ according to different segments of the larger patient population. However, another attendee suggested that you cannot expect to have a common set of preferences for any given patient population, and so the elicitation of heterogeneous preferences may not be very helpful but might make it more complicated to formulate a single recommendation for the whole population when there are clear differences in preferences. This also demonstrates that not every HTA will necessarily benefit equally from patient preferences data being available.

## 6.4 Concerns related to the use of patient preferences studies in HTA

Attendees expressed concern about the possible cost of conducting patient preferences studies and who should fund them. If patient preferences data are to be sufficiently robust and of high quality then the preferences studies are likely to be relatively complex studies to design and implement, and therefore lengthy and cost-intensive. Another attendee suggested that if insufficiently resourced it becomes increasingly likely that the studies would be flawed leading to poor quality data. However, the cost of a single patient preferences study probably will only be a fraction of the total cost of clinical development of a technology, and would be far less costly than the vast majority of clinical trials.

There were several points of contention raised in relation to the potential industry sponsorship of patient preferences studies. The first concerned impartiality of the data and ways in which significant bias into the design of these studies could be introduced. The point was also made, however, that bias could just as easily be introduced to the design through the researcher not getting the questions quite 'right' due to the complexity of some of the preference elicitation methods and the lack of 'gold standard' guidance on how to construct a study design for any given application. Furthermore, this risk could be mitigated by HTA bodies

and regulatory agencies offering scientific advice on the appropriateness of a patient preferences study's design if manufacturers wish to submit the results of a preferences study as part of regulatory or HTA processes.

Secondly, a related concern was voiced regarding 'ownership' of the data and the possibility that manufacturers might produce preference data for different treatments in the same disease area that gave contradictory weights for different attributes, thus potentially being a source of confusion for decision makers rather than making decisions easier. It was suggested that patient advocacy organisations might be well placed to take control of these studies in order to ensure a degree of impartiality and to help ensure the production of data that were relevant across the disease and internationally, rather than being too treatment-specific. This would have the added benefit of ensuring a minimum duplication of preferences studies within a disease area, and associated reduction in respondent fatigue from being asked to participate in multiple similar studies. But here again is the contentious issue of preference heterogeneity within specific populations, caused by a variety of demographic and disease trajectory-related factors, which will potentially inhibit the generalisation of preferences across the whole international population for a given disease.

## 6.5 What role do HTA bodies have?

Several workshop attendees suggested that HTA bodies such as NICE need to recognise the limitations of some of their preferred methods and that there might be situations where different types of data, including patient preferences data generated through a well-designed DCE, could aid decision makers in developing their recommendations. HTA bodies determine which types of data manufacturers have to collect and present as part of their submissions. As a result, HTA bodies, including NICE, are in a position to be highly instrumental in promoting the use of patient preferences studies as part of HTA policy.

## 7. CONCLUSIONS

---

This report summarises the research activities and main findings of a two-year project that was a collaboration between Myeloma UK and NICE. The aim of the project was to explore how methods for measuring patient preferences could be used in HTA. We performed a literature review, a survey and focus group with myeloma patients, and a multi-stakeholder workshop to answer the main research questions. We conclude that there is a clear scope for better use of quantitative patient preferences studies within HTA. The research conducted with myeloma patients demonstrated how the disease impacts their lives in many and sometimes different ways. Such research could provide the basis for a DCE that would be able to elicit patient preferences for different treatment options for myeloma. Therefore, it is important for HTA bodies, including NICE, to consider how these preferences can be captured and where in their processes they might have most value.

We identified DCEs and BWS as the most obvious study types for eliciting patient preferences as the literature review reported in chapters 3 and 5 clearly identified these as being the most methodologically robust. DCEs are already commonly used in healthcare settings whereas BWS is a newer and less familiar method. Notwithstanding, both methods appear to be able to generate valid results that allow insights in the relative importance of different treatment attributes, and the trade-offs that patients are willing to make between treatment options. But these methods are unlikely to be universally appropriate. We found that selecting the most appropriate method will depend on the specific research question, the size of the population to be studied, and resource constraints. Other methods that provide relative preference weights without trade-off data do have certain advantages that may make them appropriate alternatives when time, ease of comprehension, small sample sizes or cost are significant factors.

A final question is, where within HTA processes in general – and within NICE processes, specifically – these methods would have most value. This question cannot be completely answered without piloting the methods identified as most suitable, which was beyond the scope of this project. Notwithstanding, we conclude that there are three areas of NICE work where these methods could be of most value. First, we see a possible role for patient preferences studies in the scientific advice context, when companies are exploring the value proposition of their products but might not have mature data yet. Here, patient preferences studies could support a case for demonstrating a clear unmet medical need, a large gap between treatments that are desired by patients and what is currently available in the NHS, or as a method for selecting appropriate endpoints for clinical trials, such that endpoints targeted by technologies are aligned with the attributes that matter most to patients.

Second, patient preferences methods could be part of an evidence submission within a technology appraisal or for highly specialised technologies where they could provide a more comprehensive insight into patient preferences during committee meetings that currently rely on patient testimony. The workshop



discussions, literature review, and research with myeloma patients made it clear that not all appraisals might benefit equally from having the results of a patient preferences study available. We found that a first requirement would be that a quantitative study needs to be feasible given the sample size requirements. A second requirement would be that there was potential for the patient preferences study to have added value, for example in case of comparing two very different treatment options (operation versus drug treatment), large preference heterogeneity among patients which would render patient testimony less representative, or the treatment having an impact on important benefits to patients that would not be well captured within standard methods. Along similar lines, there are also clearly identifiable situations where patient preferences data would be not make any difference to a recommendation, such as cases where there technology was clearly very cost effective or very cost ineffective, or when comparing treatment options without distinctly different attributes.

Third, within NICE clinical guidelines these studies could also help identify what treatments along a clinical pathway might be preferred by different subgroups of patients. More work is needed to better identify, and test, whether patient preferences studies indeed would be of benefit to the work NICE does.

Although preference elicitation methods have been around for a long time, there has been limited investigation of how these methods can best be used in the context of the work of HTA bodies. As our research has been exploratory, we have not yet tested the practical implications of using these data in HTA. Therefore it remains to be seen how these methods could be implemented alongside current HTA methods in such a way that the data are of most benefit to the production of high quality guidance. Future work, needs to include pilots and/or simulation studies to better understand how these patient preferences studies would be used in practice. HTA bodies, such as NICE, are perfectly positioned to take a central role in further testing how patient preference data should be generated in a meaningful way within HTA.

## 8. REFERENCES

---

Adamowicz W, Louviere J, Williams M (1994) Combining revealed and stated preference methods for valuing environmental amenities. *Journal of environmental economics and management* 26(3), 271-92.

Ali S, and Ronaldson S (2012) Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods. [Review]. *British Medical Bulletin* 103(1), 21-44.

Anderson M, and McCleary K (2015) From passengers to co-pilots: Patient roles expand. *Science Translational Medicine* 7(291), 291fs25.

Bridges J F, and Jones C (2007) Patient-based health technology assessment: a vision of the future. *International Journal of Technology Assessment in Health Care* 23(1), 30-35.

Cheung K L, Wijnen B F, Hollin I L, Janssen E M, Bridges J F, Evers S M, Hiligsmann M (2016) Using best-worst scaling to investigate preferences in health care. *Pharmacoeconomics* 34(12), 1195-209.

Clark M D, Determann D, Petrou S, Moro D, de Bekker-Grob E W (2014) Discrete choice experiments in health economics: a review of the literature. *Pharmacoeconomics* 32(9), 883-902.

de Bekker-Grob E W, Ryan M, Gerard K (2012) Discrete choice experiments in health economics: a review of the literature. *Health Economics* 21(2), 145-172.

Bekker-Grob E W, Berlin C, Levitan B, Raza K, Christoforidi K, Cleemput I, Pelouchova J, Enzmann H, Cook N, Hansson M G (2017) Giving Patients' Preferences a Voice in Medical Treatment Life Cycle: The PREFER Public-Private Project. *The Patient* 10(3), 263-6

Devlin N J, Shah K K, Feng Y, Mulhern B, van Hout B (2018) Valuing health-related quality of life: An EQ-5D-5L value set for England. *Health Economics* 27(1), 7-22

Egbrink M O, IJzerman M J (2014) The value of quantitative patient preferences in regulatory benefit-risk assessment. *Journal of Market Access & Health Policy* 2(1):22761

European Medicines Agency (2012) Benefit-risk methodology project: work package 4 report: benefit-risk tools and processes.

[https://www.ema.europa.eu/documents/report/benefit-risk-methodology-project-work-package-4-report-benefit-risk-tools-processes\\_en.pdf](https://www.ema.europa.eu/documents/report/benefit-risk-methodology-project-work-package-4-report-benefit-risk-tools-processes_en.pdf)

(accessed September 2018)

European Medicines Agency (2013) The patient's voice in the evaluation of medicines: How patients can contribute to assessment of benefit and risk.

[http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Report/2013/10/WC500153276.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Report/2013/10/WC500153276.pdf)

(accessed January 2017)

Feeny D H, Eckstrom E, Whitlock E P, and Perdue LA (2013) A primer for systematic reviewers on the measurement of functional status and health-related quality of life in older adults. Research White Paper. Agency for Healthcare Research and Quality, Rockville, MD.

US Food and Drug Administration (2016) Patient preference information—voluntary submission, review in premarket approval applications, humanitarian device exemption applications, and de novo requests, and inclusion in decision summaries and device labeling.

<https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446680.pdf>

(accessed September 2018)

Henshall C, Schuller T (2013) Health technology assessment, value-based decision making, and innovation. *International Journal of Technology Assessment in Health Care* 29(4), 353-359.

Hensher D A (1994) Stated preference analysis of travel choices: the state of practice. *Transportation* 21(2), 107-33.

Ho M P, Gonzalez J M, Lerner H P, Neuland C Y, Whang J M, McMurry-Heath M, Hauber A B, and Irony T (2015) Incorporating patient-preference evidence into regulatory decision making. *Surgical Endoscopy* 29(10), 2984-2993.

Ho M, Saha A, McCleary K, Levitan B, Christopher S, Zandlo K, Braithwaite R S, Hauber A B, and Medical Device Innovation Consortium's Patient Centered Benefit-R (2016) A Framework for Incorporating Patient Preferences Regarding Benefits and Risks into Regulatory Assessment of Medical Technologies. *Value in Health* 19(6), 746-750.

Hughes D, Waddingham E, Mt-Isa S, Goginsky A, Chan E, Downey G F, Hallgreen C E, Hockley K S, Juhaeri J, Lieftucht A, Metcalf M A (2016) Recommendations for benefit–risk assessment methodologies and visual representations. *Pharmacoepidemiology and Drug Safety* 25(3), 251-62.

Johnson F R, and Zhou M (2016) Patient Preferences in Regulatory Benefit-Risk Assessments: A US Perspective. *Value in Health* 19(6), 741-745.

Johnson G (2017) *Pharma-Specific Conjoint: Evidence-based conjoint analysis for an evidence-based industry*. London Scientific Publishing, London.

Johnson FR, Beusterien K, Özdemir S, Wilson L (2017) Giving patients a meaningful voice in United States regulatory decision making: the role for health preference research. *The Patient* 10(4), 523-6.

Levitan B, Hauber A B, Damiano M G, Jaffe R, Christopher S (2017) The ball is in your court: agenda for research to advance the science of patient preferences in the regulatory review of medical devices in the United States. *The Patient* 10(5), 531-536

Mott D J (2018) Incorporating Quantitative Patient Preference Data into Healthcare Decision Making Processes: Is HTA Falling Behind? *The Patient* 11(3), 249–252

Mühlbacher A C, and Johnson F R (2016) Choice Experiments to Quantify Preferences for Health and Healthcare: State of the Practice. *Applied Health Economics & Health Policy* 14(3), 253-266.

Mühlbacher A C, Zweifel P, Kaczynski A, and Johnson F R (2016a) Experimental measurement of preferences in health care using best-worst scaling (BWS): theoretical and statistical issues. *Health Economics Review* 6(2), 1-12.

Mühlbacher A C, Kaczynski A, Zweifel P, Johnson F R (2016b) Experimental measurement of preferences in health and healthcare using best-worst scaling: an overview. *Health Economics Review* 6(1), 1-9.

Mühlbacher A C, Bridges J F, Bethge S, Dintsios C M, Schwalm A, Gerber-Grote A, Nübling M (2017) Preferences for antiviral therapy of chronic hepatitis C: a discrete choice experiment. *The European Journal of Health Economics* 18(2), 155-65.

Rawlins M, Barnett D, and Stevens A (2009) Pharmacoeconomics: NICE's approach to decision-making. *British Journal of Clinical Pharmacology* 70, 346-349.

Ryan M, Scott D A, Reeves C, Bate A, Van Teijlingen E R, Russell E M, Napper M, and Robb C M (2001) Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technology Assessment* 5(5), 1-186.

Sattler H, and Hensel-Borner S (2000) A comparison of conjoint measurement with self-explicated approaches, in Gustafsson A, Herrmann A, and Huber F (Eds.) *Conjoint measurement: Methods and applications*. Springer Verlag, Berlin.

Schmidt K, Babac A, Pauer F, Damm K, von der Schulenburg J M (2016) Measuring patients' priorities using the Analytic Hierarchy Process in comparison with Best-Worst-Scaling and rating cards: methodological aspects and ranking tasks. *Health economics review* 6(1), 50.

Tervonen T, Gelhorn H, Sri Bhashyam S, Poon J L, Gries K S, Rentz A, Marsh K (2017) MCDA swing weighting and discrete choice experiments for elicitation of patient benefit-risk preferences: a critical assessment. *Pharmacoepidemiology and drug safety* 26(12), 1483-91.

van Til J, and IJzerman M J (2014) Why should regulators consider using patient preferences in benefit-risk assessment? *Pharmacoeconomics* 32(1), 1-4.

Wainwright D M (2003) More 'con' than 'joint': Problems with the application of conjoint analysis to participatory healthcare decision making. *Critical Public Health* 13(4), 373-380.

Whitty J A, Walker R, Golenko X, and Ratcliffe J (2014) A think aloud study comparing the validity and acceptability of discrete choice and best worst scaling methods. *PLoS One* 9(4), e90635.

Wijnen B F, van der Putten I M, Groothuis S, de Kinderen R J, Noben C Y, Paulus A T, Ramaekers B L, Vogel G C, and Hiligsmann M (2015) Discrete-choice experiments versus rating scale exercises to evaluate the importance of attributes. *Expert Review of Pharmacoeconomics & Outcomes Research* 15(4), 721-728.

# APPENDIX 1: STUDY DESIGN AND METHODS

---

In order to answer the primary research question, two complementary research methods were employed: a survey of myeloma patients and a focus group comprising equal numbers of patients and carers.

## Survey

The survey was designed to collect a wide range of demographic and condition-specific data from patients, including age, employment and household status, length of time with the condition, number of treatment courses received and self-reported health status. Crucially, none of these data are sufficient to compromise respondents' anonymity. The survey also collected qualitative responses on several questions, including the impact myeloma has on the daily lives of themselves and their family/friends, the good outcomes that they desire, and the bad outcomes that they hope to avoid. The internet-based survey was hosted on the Snap Survey platform and the link was disseminated to as many potential respondents as possible. Eligible patients were those with myeloma in the UK; there were no exclusion criteria. The survey link was disseminated by Myeloma UK.

Recent work has explored the preferences of myeloma patients in the context of risk-benefit assessment<sup>3</sup>. But previous surveys with this patient group have not used open-ended questions that allow respondents freedom to identify all potential attributes of importance to them, and instead have focussed on refining and weighting pre-defined attributes.

## Focus group

To complement the results of the survey, a focus group was conducted with 10 respondents, half of these being myeloma patients, and half carers/relatives of people with myeloma. This was a convenience sample from a local myeloma support group. The focus group lasted for 1.5 hours and was structured to achieve two principal objectives.

Firstly, coded themes from the qualitative questions in the survey were presented to the group and formed the foundation of discussion on certain themes, as determined by the group. This was in part to examine the face validity of the themes identified in the survey, and also to elicit greater contextual analysis

---

<sup>3</sup> Douwe Postmus et al. (2018) "Individual Trade-Offs Between Possible Benefits and Risks of Cancer Treatments: Results from a Stated Preference Study with Patients with Multiple Myeloma", *Oncologist*, 23(1): 44-51.

from the group on these themes. Survey data was presented simply, using lists of identified themes in order to highlight the breadth of the issues described in respondents' qualitative answers. Coding frequencies (although not necessarily an indicator of relative importance) were only revealed to the group at the end of the discussion so as to prevent biasing the responses. Secondly, following a brief discussion about the range of available methods for measuring patient preferences and some of their relative strengths and weaknesses (utilising summary tables of characteristics developed for this purpose) discussion was guided towards how well these might be suited to measuring the kinds of outcomes and other preferences identified, and what might be the defining criteria that could inform the selection of a particular preference elicitation method for this specific patient population, such as the relative cognitive burden of different methods, for example. The focus group was recorded and transcribed verbatim.

## Data analysis

Both the quantitative and qualitative data were analysed in Excel. The survey collected patient demographic and condition-related data to ensure that respondents were representative of the wider patient population and to enable stratified analysis. Qualitative data in the survey and from the focus group were subjected to thematic content analysis<sup>4</sup>. Identification of initial themes were informed by the available literature on patient experiences of myeloma and available treatments. These themes were further developed by inductive content analysis of the data in an iterative process. This process began with an expansive approach in which the textual responses were coded using multiple, highly specific, themes. Code lists were then scrutinised, and where overlapping themes were identified these were combined into higher level themes as necessary to provide the final coding structure. Simple summary tables and graphs were used to characterise the survey results.

---

<sup>4</sup> Elo, Satu, and Helvi Kyngäs. "The qualitative content analysis process." *Journal of advanced nursing* 62.1 (2008): 107-115.

## APPENDIX 2: BEST-WORST SCALING (BWS) DESIGNS

Least important	Attribute	Most important
	Side effects	•
•	Treatment frequency	
	Treatment mode of delivery	
	Location of treatment delivery	
	Duration of treatment	

An example of a fictional BWS question - attribute case (type 1)

Least important	Attribute-level	Most important
	Side effects: Severe	•
	Treatment frequency: Weekly	
	Treatment mode: IV injection	
•	Location of treatment: Hospital	
	Duration of treatment: 10 weeks	

An example of a fictional BWS question - profile case (type 2)

Attributes	Treatment A	Treatment B	Treatment C
Side effects	Severe	Moderate	Moderate
Treatment frequency	Monthly	Monthly	Daily
Treatment mode	Oral	Injection	Oral
Treatment location	Home	Hospital	Home
Treatment duration	24 weeks	36 weeks	40 weeks
Best treatment	•		
Worst treatment		•	

An example of a fictional BWS question - multiprofile case (type 3)









Published June 2019

---

Myeloma UK is the only organisation in the UK dealing exclusively with myeloma. Our ultimate goal is to find a cure.

We are dedicated to myeloma patients – making sure they get access to the right treatment at the right time. We continually strive to improve standards of treatment and care through research, education and raising awareness of myeloma.

For more information about Myeloma UK, including details of our latest research and analysis, please visit [myeloma.org.uk](http://myeloma.org.uk).

---

**Myeloma UK**  
22 Logie Mill  
Beaverbank Business Park  
Edinburgh EH7 4HG

t: **0131 557 3332**  
e: [myelomauk@myeloma.org.uk](mailto:myelomauk@myeloma.org.uk)

[myeloma.org.uk](http://myeloma.org.uk)



Registered Charity No: SC026116