Understanding public preferences for avoiding QALY losses caused by lapses in healthcare safety and patient lifestyle choices.

Jeshika Singh[1, 2], Amanda Baine[2], Louise Longworth[2], Joanne Lord[2], Shepley Orr[3, 4] & Martin Buxton[1, 2]

[1]Multidisciplinary Assessment of Technology Centre for Healthcare (MATCH), Department of Information Systems and Computing, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK
[2]Health Economics Research Group (HERG), Brunel University, Uxbridge, Middlesex, UB8 3PH, UK
[3]Department of Civil, Environmental and Geomatic Engineering, University College London, London, WC1E 6BT, UK

1. Introduction

Economic evaluations of health interventions often use an incremental cost-per-QALY approach, particularly in the UK where it is the preferred analytic framework of the National Institute of Health and Clinical Excellence. Such analyses usually assume that 'a QALY is a QALY'; that each unit of health outcome is worth the same no matter who gains or loses it, or under what circumstances it is gained or lost. This assumption is consistent with a policy objective of health maximisation. However, in recent proposals for changing the system for pricing branded medications in the UK, the Department of Health has announced plans to develop weights that reflect the value that society places on different types of health benefits. There is some empirical evidence that the public want resource allocation decisions to be informed by considerations other than just efficiency (1,2). One area that the public may choose to prioritise is the safety of health care interventions as the risk of healthcare harming a patient may be seen as particularly undesirable.

Studies of transport safety have found that people have a particular 'dread' for public transport accidents and that they would be willing to pay a premium to avoid them (3-6). Similar intuition could be applied to healthcare safety, which might involve value judgements beyond the probability and consequence of the healthcare risk. Conversely, people may give less priority to health loss where the individual was responsible (7-9). These preferences could potentially be captured using QALY weights.

In this study we define a 'healthcare safety incident' as a situation where harm is done (or there is potential for harm) which is to some extent caused by the actions or inactions of healthcare providers or the healthcare system. We further define a 'healthcare safety intervention' as actions or policies which are designed to prevent the occurrence of safety incidents, or to prevent or mitigate the harm arising from them. This contrasts with other types of healthcare interventions which are directed at preventing or mitigating harm not caused by healthcare staff or systems; for example, interventions for lifestyle-related diseases for which the individual patient is perceived to be responsible or interventions for genetic diseases for which nobody is responsible.

Previous research has compared relative preferences on the broad basis of responsibility (9,10). We conducted a study to understand public preferences regarding responsibility in more detail, comparing interventions to avoid QALY losses in a number of different contexts.

2. Method

Overview of Valuation Method

A relative valuation approach was used to compare six different contexts. Three of these represented lapses in healthcare safety: hospital associated infections, medication errors and injury to NHS staff at their workplace. The other contexts, chosen for comparison, included two to prevent harms for which the patient might be said to be wholly or partly responsible (lifestyle related diseases and sports injuries) and finally, a context for which responsibility could not be attributed to either the individual or the healthcare system (genetic disorders).

The relative valuations placed on the six contexts were captured using a person trade off (PTO) method (11,12). This allows respondents to trade off one risk against another by varying the number of people affected by the two risks until the point at which the respondent is indifferent. The value yielded using PTO is different to that obtained from methods designed to elicit health related quality of life or 'utility' (such as the standard gamble or time trade off), because it adopts an explicitly societal perspective rather than an individual choice perspective. Respondents to PTO questions are more likely to express "non-maximizing principles" and may incorporate distributive preferences (3-6). This method has been used in several transport

studies, where it is known as the 'matching' technique, and it has also been used recently in the Social Value of a QALY project (13).

Questionnaire Design

There are 15 possible pairings of the six contexts. The service to prevent harm from genetic disorders was used as the baseline for comparison, as this is the most neutral in attribution of responsibility. Each respondent was asked six PTO questions; the first five being against the baseline of genetic disorder, and the sixth being one of the ten remaining comparisons (randomly selected for participating individuals). To ensure that the respondents understood the trade-off process, a simple analogy and a practice question and answer were provided. The contexts were described using the examples in Table 1.

Table 1: Description of contexts presented to respondents

	Services	**Risk Examples**	**Service Examples**	**Responsibility**	**Who is affected?**
1	Service to prevent lifestyle related diseases. **(Lifestyle Disease)**	Diseases caused by smoking, drinking too much, or not exercising	Stop smoking program; healthy eating campaign; adverts to reduce binge drinking	Patient	Patient
2	Service to prevent sports injuries. **(Sports Injury)**	Back injury while gym training, knee injury because of repeated running	Safer sports equipment; sports injury clinic	Patient	Patient
3	Service to prevent hospital associated infections. **(Hospital Infection)**	MRSA (superbug), wound infection after surgery	Better cleaning; isolation procedures once detected	NHS	Patient
4	Service to prevent patients being given incorrect drugs. **(Medication Error)**	People being given the wrong drug, or the wrong dose	Strict protocols and cross checking; Computerised prescribing and decision support	NHS	Patient
5	Service to prevent diseases due to genetic disorders. **(Genetic Disorder)**	Hereditary high cholesterol or blood pressure; inherited eczema; asthma	Early screening to find people at risk; monitoring or early treatment to reduce risk;	Nobody	Patient
6	Service to prevent injury to NHS staff at workplace. **(Staff Injury)**	Needle stick (sharps injury), back injury	Devices with built-in safety features; better lifting equipment.	NHS	NHS staff

In each PTO question, respondents were asked to choose between two preventative services (A and B) that would benefit 1,000 people each, and the health gain per person and costs to the NHS for both programmes were said to be identical (see figure 1). This encouraged respondents to trade off one context against another purely in terms of persons affected. Respondents were given the option to indicate no preference between the two services.

Figure 1: Example of PTO question used.

Setting: Please imagine that the local health service has some extra money to spend on preventive health care within your area and there are two services it could choose from. It can only afford to fund one of the services, and it cannot provide a mix of both. The services cost the NHS the same amount of money, and they provide the same health improvement.

Q1a. Which of the following two services (i.e. A or B) would you choose or is there no difference?

	OPTION A	**OPTION B**
	Service to prevent lifestyle related diseases.	**Service to prevent diseases due to genetic disorder.**
Number of people who benefit	**1000**	**1000**
Total cost to the NHS	**£200,000**	**£200,000**
Health benefit per person	**Avoid 3 months of moderate ill health***	**Avoid 3 months of moderate ill health***

*Moderate ill health is when a person has some problems walking about, some problems washing or dressing, some problems with performing their usual activities (e.g. work, study, or looking after children), moderate pain or discomfort and moderate anxiety or depression

Select **A. Prioritise service to prevent sports injuries.** ☐ **B. Prioritise service to prevent diseases due to genetic disorder.** ☐ **U. No Difference** ☐

If participants were indifferent, it was assumed that they gave both services equal value. Where the participant indicated a preference for option A or B, their strength of preference was measured by varying the number of people using a 'payment card' approach as shown in Figure 2. The participants also had the option of specifying their own value, or of indicating that the NHS should never prioritise the less-preferred service A over B.

Figure 2: Payment ladder (part of Q1a depicted in figure 1)

Q1c. When the service prevented 1000 cases each, you chose Option B. Suppose the number of people who benefit from the two services is different. How many people would Option A have to benefit so that you would change your answer and choose Option A?

Please proceed down the table and select the level at which you would change your answer to Option (not chosen in 1a). However, if you do not agree with these levels you can insert your own level. Only one answer is required here.

OPTION B **Service to prevent diseases due to genetic disorder**	**OPTION A** **Service to prevent lifestyle related diseases.**	**Prioritise A**
Prevents 1000 cases	Prevents 1100 cases	
Prevents 1000 cases	Prevents 1200 cases	
Prevents 1000 cases	Prevents 1300 cases	
Prevents 1000 cases	Prevents 1400 cases	
Prevents 1000 cases	Prevents 1500 cases	
Prevents 1000 cases	Prevents 1700 cases	
Prevents 1000 cases	Prevents 2000 cases	
Prevents 1000 cases	Prevents 3000 cases	
Prevents 1000 cases	Prevents 4000 cases	
Prevents 1000 cases	Prevents 5000 cases	
Prevents 1000 cases	Prevents 10000 cases	
Prevents 1000 cases	Prevents other (specify)	……..
NHS should not prioritise Option A over Option B		

The main survey was conducted in September 2010 using a self-completion questionnaire administered over the internet. The online survey was created in partnership with a professional market research company and administered via a large online panel. A quota system was used to determine eligibility of respondents, ensuring that the sample reflected the UK general population in terms of gender, age, geographical location and occupation (2001 UK Census). Prior to the survey, 13 cognitive interviews were conducted and based on this

feedback the payment card method was chosen. Then a pilot study (N=100) was conducted to test the range of values and intervals used on the payment card.

Analysing PTO responses

To compare individuals' relative strength of preference across different contexts, PTO ratios were calculated compared with the basecase of preventing one case of harm from genetic disorder. For example, if the respondent indicates that prevention of harm from genetic disorders for 1000 people is of equal priority to preventing harm from lifestyle related diseases for 1000 individuals, then the two services are valued as equal, and the PTO ratio has a value of 1. When respondents indicate a preference for the base case, the PTO ratio will be less than 1. For example: if they rate prevention for 10,000 cases of lifestyle related disease as equivalent to prevention for 1,000 cases of genetic disorder, then the PTO ratio is 0.1 (1,000/10,000). Conversely, when respondents indicate a preference for the comparator over the base case, the PTO ratio will be greater than 1. For example: if they rate prevention for 10,000 genetic cases as equivalent to prevention for 1,000 lifestyle cases, then the PTO ratio is 10 (10,000/1,000).

The consistency (transitivity) of preferences was tested by comparing PTO ratios calculated from the direct comparison of two non-base case contexts (the sixth PTO question that each respondent was asked) with the equivalent indirect (or 'chained') PTO ratio calculated from the PTO ratios for each base case comparison. A Wilcoxon 'sign rank' test was used to compare matched pairs of direct and indirect ratios.

Aggregating the ratios

There is no single best measure of central tendency for aggregating PTO ratios across individuals. The most common approach is to take the arithmetic mean of PTO ratios. But this violates the condition of symmetry and produces inconsistent results. Two alternative measures have been recommended: 1) taking the median of ratios derived from individuals and 2) taking the arithmetic mean of 'context ratios' (13).

Context ratios are obtained by first identifying the largest PTO ratio across the six contexts for each individual, and then dividing the individual's PTO ratios by this maximum (5). Each respondent's highest ranked context is therefore assigned a value of 1 and their lower-ranked

contexts are assigned values less than 1. Weights across individuals are then calculated by taking the mean for each of the six context ratios, and dividing by the mean context ratio for the basecase (the 'ratio of means'). This gives a relative valuation of each context compared with the basecase.

Non trade responses

In addition to allowing equal weights to contexts, the study allowed non trade responses. According to the stated preference literature, there are two ways of dealing with these lexicographic responses (14). The first is to exclude them as they represent subjective non-numeric responses. The second is to allocate non trade responses an arbitrarily large numeric response and to test this in sensitivity analysis. We used this latter approach, since the participants were allowed non trade responses in our survey.

Multivariate analysis

Apart from socio-demographic questions, respondents were asked about their attitudes to, and use of, the NHS. We included three factors that have been shown to influence risk valuations (15,16): i) the participants' perceptions of their own or their families' exposure to each of the six types of risk ('exposure'); ii) their experience of each type of risk ('familiarity'); and iii) their perception of the severity of the consequences of each risk ('dread'). Finally, information was collected on socioeconomic factors such as income, newspaper and education.

Multivariate analysis was used to investigate whether any of these individual characteristics and attitudes were associated with preferences. Separate equations were fitted for each of the five comparisons against the genetic basecase.

We tested different ways of coding the dependent variable to reflect the direction and strength of preferences. PTO ratios are bounded at zero and are not normally distributed. Taking the log of the PTO ratios improved the OLS fit; however the resulting error terms were still not normal. We therefore used an ordered probit approach, categorising preferences in three levels: 1) preference for non-genetic, 2) indifferent and 3) preference for genetic. We also tested a five-level ordered dependent variable, differentiating between traded preferences (where the respondent was willing to trade-off between the contexts at some PTO equivalence

number) from non-trade preferences. However, the assumption of proportionality of odds across response categories was not met with the five-level models.

The models were tested with a range of explanatory variables: demographic factors (age and gender), socioeconomic factors (occupational class, income category, educational qualifications, newspaper), NHS factors (frequency of use, satisfaction and whether they or a family member worked for the NHS), and attitudes to the risk category (exposure, familiarity and dread).

Qualitative analysis

For each PTO question, respondents were invited to explain their choices in a free text box. These comments were examined using thematic analysis (17).

A step-by-step process was used for this thematic content analysis and involved three distinct stages (18). The first stage was familiarisation, where the aim was to obtain an overview of some of the broader themes (19). In the second stage, coding took place and general themes were named and defined. Computer software was used to aid the coding (Microsoft Excel 2003) and to filter, categorise and graph the responses. The codes were created by one of the authors (AB) and a random sample of 100 comments was coded in parallel by another member of the team (JS). The coding differences were resolved through discussion. The third stage involved grouping the themes to form an overall coding framework.

3. Results

Preferences across pair wise comparisons

Figure 2 shows the distribution of PTO responses for the five comparisons with baseline (intervention to prevent genetic disease). The column in the middle represents the indifference responses, where both services were given equal value. For example, 405 people valued a service preventing harm from genetic disorder for 1000 people the same as a service preventing harm from hospital associated infection for 1000 people. Indifferent responses ranged from 12% to 55% of the choices. However, only 34 respondents (3%) were always indifferent across all six comparison questions that they were asked. The columns to the left of the indifference column indicate that the genetic context was preferred and those to the right

indicate that the other context was preferred. The column on the far left of Figure 2 shows the numbers of non-trade responses where the genetic disorder context was preferred. And similarly, the column on the far right shows the numbers of non-trade responses where the other context was preferred. 59% of respondents gave a non trade response for at least one of the six PTO questions that they were asked.

The median PTO ratios and context ratios estimated from the five baseline comparisons are shown in Table 2. The non trade responses were included in these summary scores by attaching a large value (100,000). Sensitivity analysis showed that the median and context ratio values did not change beyond this value. Avoiding harms from hospital infections, medication errors and genetic disorders were given an equal median weight. But relative priority using the context ratio gave greater priority to hospital infection (1.31) than interventions for genetic disorders (1.00). This difference can be understood with reference to the distribution of PTO preferences in Figure 1. Although 405 respondents (39%) were indifferent between the hospital infection and genetic disorder contexts, 429 (42%) expressed a preference for hospital infection, and of these 162 (38%) said that the NHS should always prioritise prevention of harm from hospital infections.

The other relative valuations calculated by the context ratio method were similar to those calculated by the median method: medication errors were given very similar weight to genetic disorders (1.07); whereas staff injury, lifestyle disease and sports injuries were given less weight (0.71, 0.65 and 0.41 respectively).

Figure 2: Preference across pair wise comparisons

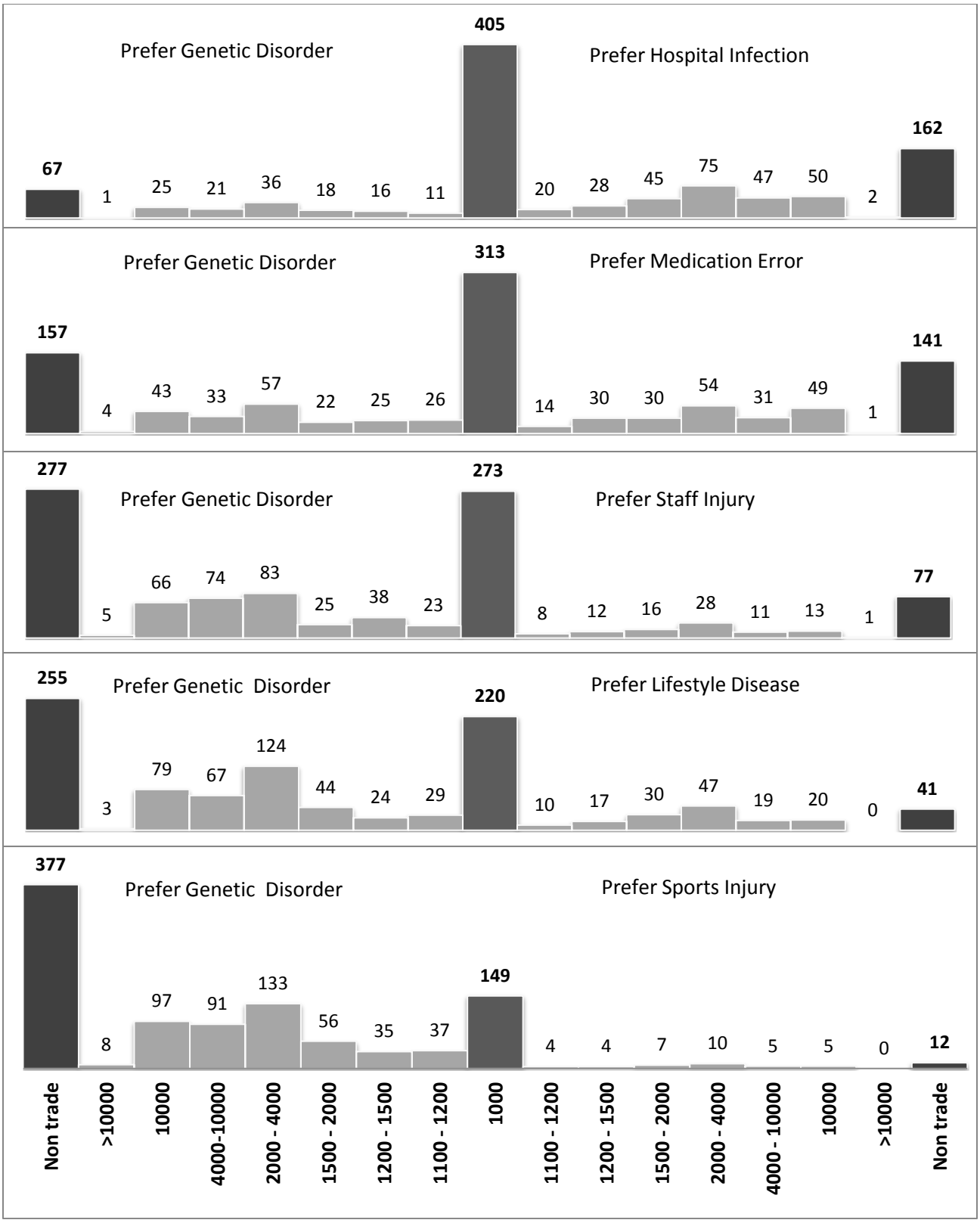


Table 2: Person trade-off valuation results

	Median[2]	Context Ratio Mean[2]	Relative Priority[1]
Hospital Infection	1.00	0.73	1.31
Medication Error	1.00	0.59	1.07
Genetic Disorder	1.00	0.55	1.00
Staff Injury	0.67	0.39	0.71
Lifestyle Disease	0.50	0.36	0.65
Sport Injury	0.20	0.23	0.41
[1]Mean of the context compared with mean for basecase context (genetic disorder) *[2]Calculated by attaching a person trade off value of 100,000 to non-trade responses*			

Consistency of preferences

The Wilcoxon "sign rank" test showed that there was no statistically significant difference between individuals' direct and chained estimates, suggesting that the responses are internally consistent (z = -1.262).

Results of the ordered probit models are displayed in Table 3. Different factors were found to be statistically significant for preferences in the five comparisons. Positive coefficients indicate a greater latent preference for the basecase (genetic) context relative to the reference category. Negative coefficients imply a greater preference for the non-base case comparator (non-genetic) relative to the reference category. Model specifications were tested using the RESET test, which rejected the hypothesis that the model was mis-specified.

Older people (aged 60 and over) were significantly more likely to favour interventions for hospital infections over interventions for genetic disorders.

Table 3: Ordered probit model of direct preference

Independent Variable (Reduced model)	N	Lifestyle diseases	Sports injuries	Hospital infections	Medication errors	Staff injuries
Sex						
Male	506		-0.207**			
Female(reference category)	524					
Age						
18-29 (reference category)	162					
30-39	213			0.086	0.114	0.291**
40-49	151			0.125	0.329**	0.159
50-59	214			0.029	0.292**	0.206*
60+	290			-0.217**	0.0967	-0.060
NHS Use						
More than once a month (reference category)	85					
Once a month	233	0.260*				
Once in 3 months	284	0.260*				
Twice a year	182	0.291*				
Once a year or less	230	0.070				
Don't know	16	-0.482				
Perceptions of risk						
***Fear (with 1 being highest score and 6 being lowest score*)**						
1 (reference category)						
2		0.334**		0.314***	0.139	0.410
3		0.309***		0.617***	0.573***	0.540
4		0.562***		0.734***	0.645***	0.787
5		0.599***		0.786***	0.966***	1.128**
6		0.905***		0.714	0.829**	1.122**
Higher risk than others						
Yes (reference category)						
No			0.361***			
Experienced the risk						
Yes (reference category)						
No		0.224**				
Model Diagnostics						
Cut 1		-0.154	-1.522	0.058	0.043	0.162
Cut 2		0.518	-0.693	1.207	0.863	0.989
Log likelihood		-942.172	-600.754	-1036.108	-1091.195	-974.724
Pseudo R^2		*0.027*	*0.014*	*0.040*	*0.034*	*0.019*
Specification	***Chi squared***	*6.03*	*5.59*	*10.86*	*12.85*	*8.54*
	(p-value)	*(0.871)*	*(0.061)*	*(0.285)*	*(0.170)*	*(0.481)*

Lower levels of dread for each of the three safety contexts were associated with a lower probability of favouring them over the genetic comparator. People were more likely to favour

interventions for lifestyle diseases if they reported higher levels of dread, more frequent use of the NHS, or personal or family experience with lifestyle-related diseases. They were also more likely to favour interventions for sports injuries when they reported that they were at higher than average risk of such injuries. Also men were more likely to favour services for sports injuries than women.

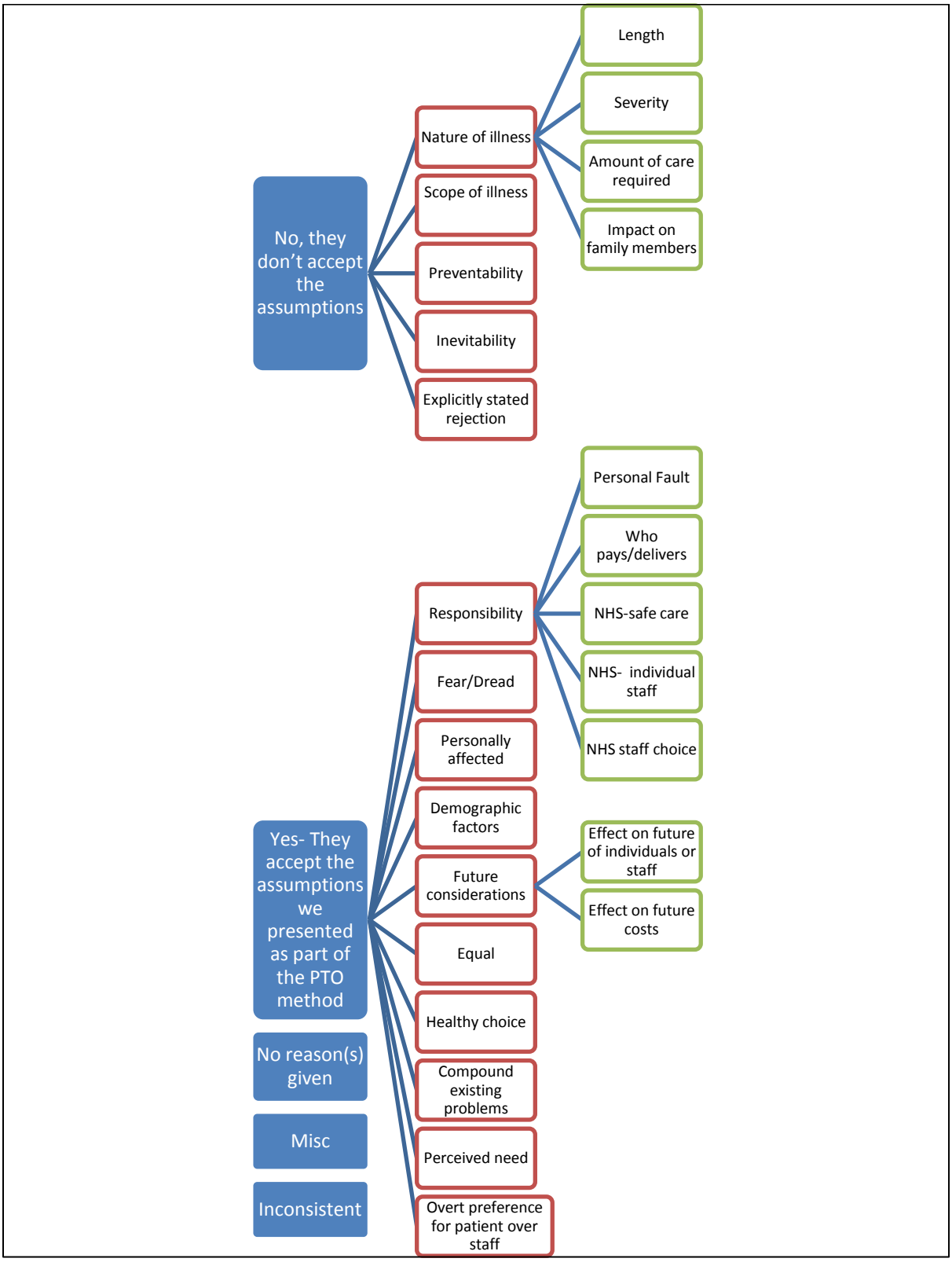
Qualitative data

Respondents had the option of ticking 'no comment', but most chose to explain their preferences. There were 3,743 comments overall but only 3,286 comments were included in the final analysis after excluding comments which were inconsistent or did not provide explanatory information. Examples of such comments include: "*hi*", "*unable to doanic disease genetthing regarding*", "*it just sounds right*" and "*Because that [sic] the way I believe it should be*". Multiple subjects/codes were contained with some comments, and a total of 4050 excerpts were identified. A total of 23 themes or codes emerged upon detailed analysis.

The parallel coding exercise led to a reliability rate of 88%. The remaining discrepancies were minor and often involved one reviewer adding a second code where the first codes matched. Differences were resolved through discussion and ultimately full agreement was reached. A coding-framework was used to categorise the themes (Figure 3).

Two primary categories were identified based on whether or not the respondents' comments rejected stated assumptions in the PTO questions. Respondents were told at the start of each PTO question that equal numbers of people would be helped by any of the interventions that the cost was equal, and that patients would be in a state of moderate ill health for three months without the intervention regardless of which intervention was specified. Themes which questioned the assumptions were related to the nature of the illness, the scope or prevalence of the illness, the preventability or inevitability, or responses explicitly rejecting the assumption. All other themes were put in second group where the responses did not disagree with the assumptions of the PTO exercise outright (it is not possible to infer if the second group accepted the assumptions), but were instead giving additional reasons for their choice.

Figure 3: Coding framework



Acceptance of PTO assumptions

Most (81%) of coded excerpts did not challenge the assumptions of the PTO assignment. But 19% directly or indirectly contradicted the stated assumptions. The most common contradiction related to the nature of the illness with respondents stating that they believed one type of condition would be more severe or differ in the amount of care required, despite being told at the start of each PTO question that all patients would experience 3 months in a moderate health state without the intervention. An example of these comments is: "*GENETIC DISEASES ARE GENERALLY MUCH MORE SERIOUS THAN SPORTS INJURIES...[sic]".* Prevalence or likelihood of illness was the next most cited reason. Another 50 excerpts explicitly rejected the equal cost and consequence assumptions. For example: *"Avoidance of incorrect drug administration should not cost anywhere near that of treatment so less needs to be spent relatively".*

Differences within same responsibility

Some respondents differentiated within the same category of responsibility. For example, personal responsibility dominated as the reason why respondents preferred genetics over both lifestyle illnesses and sports injuries. However a small number of participants preferred sports injuries over genetic disorders because of the "nature of illness", especially severity and length of illness. Otherwise the explanations given were fairly homogeneous across the two contexts. Upon direct comparison of the two contexts, respondents valued them differently as some thought sports injuries were less severe, while others thought lifestyle diseases were more prevalent. And a few made a healthy choice distinction for sports injuring such as "*people do sports to remain healthy so more should be done for these people than people who smoke etc.*.[sic]".

In contexts related to NHS responsibility, most comments indicated that it was not the patient's fault but also that the NHS had a responsibility to provide clean, safe and effective care. Many respondents also felt that hospital acquired infections already affected/would affect more people than genetic disorders ("*Because the risk of becoming infected in hospital will affect more people than genetic disorders*").

The NHS Staff Injury context was the least preferred among the three NHS responsibility contexts. While many different codes were cited to explain this, the top three were:

responsibility (where respondents felt that it was either the NHS' responsibility to provide adequate health and safety training, that it was the individual staff members responsibility to take the necessary precautions, or that it was the staff members' choice to have gone into this line of work), an overt preference for patients over staff (e.g. "*the NHS is there to provide a service to the public, not to its staff*") and concern for future consideration which related to the potential for a reduction in the NHS' ability to care for patients if NHS staff are injured (e.g. "*we have to ensure nhs [sic] staff are protected otherwise there would be no staff left or wanting to work to look after us*").

Factors taken into consideration

Overall a little more than half of the responses cited a responsibility theme (2151) as the main reason for their preference. This is not surprising since the PTO study was designed to vary contexts based on responsibility. The remaining coded excerpts referred to something besides responsibility. The next most prominent theme was equality or fairness: "*Both provide the same results*" and "*I think both are equally important and the drastic improvement in both would make a huge difference to people's experience of the nhs and the way society functions..[sic]*" Followed by future considerations (*"More people would eventually benefit with lifestyle choices as a new healthy attitude would be passed on to future generations not just those directly targeted*") and personally affected/personal experience ("*My grandfather passed away due to MRSA caught in hospital, I can see the devastating effects and also the extra resources and strain this puts on an already strained nhs.. [sic]*").

Some of the factors taken into consideration varied according to context. For example some preferred hospital associated infection to avoid compounding existing problems. They reasoned that if someone comes to hospital in the first place, it's because they need medical attention, and that catching a new infection while in the hospital would only aggravate their health problems. Also the theme of fear or 'dread' only came up with hospital associated infections.

4. Discussion

The results suggest that members of the public favour QALYs gained in some contexts over others, deviating from the 'QALY is a QALY' paradigm. Only 3% of respondents were always indifferent between QALY loss prevented in different contexts. There was no clear and consistent pattern of preferences for the safety interventions however. The balance of opinion favoured prevention of QALY loss from hospital associated infections compared with preventing equivalent QALY loss from genetic disorders. But views were mixed for services to avoid medication errors, and most respondents gave lower priority to prevention of harm from workplace injuries to healthcare staff. Due to the very different patterns of preferences across the three safety contexts, we chose not to summarise them in a single score.

The reasons for these context effects were investigated through qualitative data analysis. Some comments suggested that the assumptions in the PTO questions were not accepted, but most respondents did not overtly question these assumptions. Responsibility was cited most often as the main differentiating factor, but other factors were mentioned such as fairness/equality, future consideration and personal experience.

We also found some quantitative evidence in our multivariate analysis that preferences may be related to individuals' perceptions of their exposure to the different types of risk and dread of their consequences, as well as (possibly) other personal characteristics. However, these factors only explained a small part of the wide variations in stated preferences that we observed. The methods of multivariate analysis were limited however, because the PTO equivalence values were censored and heavily skewed. This made it difficult to incorporate strength of preference in the dependent variable.

Our decision to use the PTO was motivated by our desire to compare valuations of safety interventions across the different contexts using a direct and transparent method that could be carried out over a large sample size (21-23). As Green (2001) has argued, PTO asks the right questions; inviting people to make direct policy-relevant choices. A key element of this policy relevance is that PTO questions are directed at eliciting social valuations on behalf of a community, rather than personal preferences based in individual and family interests. However, it is clearly impossible to entirely disentangle social and personal preferences. Our

regression analysis showed that perceptions of personal risk such as dread, exposure and experience were associated with preferences, although these associations were weak.

There is no consensus in the literature about how to aggregate PTO results across individuals. One question is how to deal with non trade or dominant responses, which are common in stated preference surveys (14,24,25). Excluding non trade responses may result in removal of valid preferences causing sample selection bias as well as reducing sample size; hence they were included in our results by attaching large numerical responses and conducting sensitivity analysis. A second question in aggregating PTO responses is what measure of central tendency to use. This is not just a technical issue, but a normative one. Although this issue arises for any preference elicitation technique (including for example time trade off or willingness to pay methods), particular difficulties arise for the PTO as there is a tendency to find extreme responses to PTO questions (very high equivalence values or non trade responses to which high values are imputed). These outliers would have a big effect on the arithmetic mean of the PTO ratios, which may be seen as unfair since by expressing extreme preferences a minority of people can have a disproportionate influence on the overall result.

The context ratio approach dampens the impact of outliers by first normalising each individual's PTO ratios so that their highest ratio has a value of 1, and scaling all other contexts relative to this anchor. The means of these rescaled ratios are then taken for each context, and the final relative valuations computed with respect to a common basecase - genetic disorder in our case. The effect of this process on the relative weight given to each individual in the final scores is unclear. We therefore also present our results using the simpler median of PTO ratios. This avoids the influence of outliers by finding the PTO ratio for which equal numbers of respondents give a higher and a lower value, but consequently no account is taken of peoples' strength of preference above or below this value. There are two main differences in our results obtained by these methods: the context ratio approach gives greater relative priority to hospital infection (1.31) than the median approach (1.00); and the context ratio approach gives greater priority to sports injury (0.41) than the median approach (0.2).

The survey was conducted on the web using an online panel. Concerns have been raised about the level of engagement of participants, their understanding of the tasks and the

representativeness of samples in online surveys. However, evidence suggests that PTO measurements using computer elicitation produce results of similar quality to those from face-to-face interview (20). Our study used a large sample, recruited so as to reflect the socio-demographic characteristics of the general population. Whilst there may still be differences when compared to the general population, surveys administered using other techniques may also be prone to problems of generalisability. The fact that so many respondents (73%) took the time to write comments in our survey also suggests that they were very engaged.

Our results are consistent with previous research in this area which has found that participants tend to prefer scenarios that benefited patients whose illnesses were caused by healthcare error over scenarios where patients' lifestyles were a contributing factor to their illnesses and over scenarios where illnesses were not caused by healthcare error or patient lifestyle (26). Another recent study reported that the public give higher priority to interventions for diseases where the patient has no control over the cause of the disease such as inherited disease and children's illness, and lower priority to programs for illnesses which were 'self inflicted'(27). Another study which explicitly examined preferences for treatment of liver diseases found that lower priority was given to patients who bore some responsibility for their illness (8). These results are comparable to our findings, suggesting that culpability matters. However our findings further suggest that people differentiate between contexts with the same causation and that the specifics of the context matters.

5. Conclusion

This study has produced new evidence on public priorities attached to interventions to improve healthcare safety, compared with other types of healthcare intervention. We elicited preferences for three healthcare safety contexts (services for hospital infections, medication errors and staff injuries) and three non-safety contexts (services for genetic disorders, lifestyle-related diseases and sports injuries). Despite being told that the interventions were equivalent in terms of cost, health gain and patient characteristics, patterns of preferences differed substantively within the three safety contexts, as well as between the safety and non-safety contexts. This suggests that there may not be a simple premium attached to healthcare safety, but that priorities differ between contexts in a more nuanced way. In the absence of answers to

these questions, we consider that it would be premature to use the results of this study to provide direct QALY weights for policy evaluations or to inform policy decisions more informally. It remains debateable whether such preferences if properly understood should inform policy or not, but in either case the evidence needs to be robust.

Suggestions for discussion

- Ways of conducting multivariate analysis with dependent variable being strength of preference rather than ordered category
- Both measures of central tendency, median and context ratio has obvious limitations. Any ideas on summarising the PTO ratios?
- Are there other ways of conducting qualitative analysis to capture the richness of the data?

REFERENCES

(1) Cookson R, Drummond MF, Weatherly H. Explicit incorporation of equity considerations into economic evaluation of public health interventions. Health Economics, Policy and Law 2009; 4:231–245.

(2) Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. Health Economics 2005; 14:197-208.

(3) Jones-Lee M, Loomes G. Scale and context effects in the valuation of transport safety. Journal of Risk and Uncertainty 1995; 11:183-203.

(4) Jones-Lee M, Metcalf H, Chilton S, Loomes G, Robinson A, Covey J, et al. Valuation of health and safety benefits: Dread risks. 2007;RR541.

(5) Chilton S, Covey J, Hopkins L, Jones-Lee M, Loomes G, Pidgeon N, et al. Public perceptions of risk and preference based values of safety. The Journal of Risk and Uncertainty 2002;25(3).

(6) Covey J, Robinson A, Jones-Lee M, Loomes G. Responsibility, scale and the valuation of rail safety. Journal of Risk and Uncertainty 2010;40:85-108.

(7) Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs. egalitarianism – an Australian survey of health issues. Social Sci Med 1995;41(10):1429-1437.

(8) Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. Health Economics 2000;9:137-48.

(9) Edlin R, Tsuchiya A, Dolan P. Public preferences for responsibility versus public preferences for reducing inequalities. Health Economics 2011;DOI: 10.1002/hec.1799.

(10) Dolan P, Tsuchiya A. The social welfare function and individual responsibility: Some theoretical issues and empirical evidence. Journal of Health Economics 2009;28:210-220.

(11) Nord E. The person trade-off approach to valuing health care programs. Medical Decision Making 1995;16:201-8.

(12) Nord E. Cost-Value Analysis in Health Care: Making Sense out of QALYs. : Cambridge University Press; 1999.

(13) Baker R, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G, et al. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. Health Technology Assessment 2010;14(27).

(14) Damschroder LJ, Roberts TR, Zikmund-Fisher BJ, Ubel PA. Why people refuse to make tradeoffs in person tradeoff elicitations: a matter of perspective? Medical Decision Making 2007;27:266.

(15) Slovic P. Perception of risk. Science 1987;236:280-6.

(16) Slovic P, Fishchhoff B, Lichstenstein S. Facts and Fears: Understanding Perceived Risk. In: Slovic P, editor. The Perception of Risk: Earthscan; 2000. p. 137.

(17) Green J, Thorogood N editors. Qualitative methods for health research. : London:SAGE Publications; 2009.

(18) Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006:77-101.

(19) Bryman A, Burgess RG editors. Analyzing Qualitative Data. : Abingdon: Routledge; 1994.

(20) Damschroder LJ, Baron J, Hershey JC, Asch DA, Jepson C, Ubel PA. The validity of person tradeoff mesaurements: randomized trial of computer elicitation versus face to face interview. MDM 2004;24:170.

(21) Dolan P. The measurement of individual utility and social welfare. Journal of Health Economics 1998;17:39-52.

(22) Nord E, Richardson J. The measurement of perspective in the measurement of quality adjusted life years. Medical Decision Making 1997;17:33-4.

(23) Schwappach DLB. The equivalence of numbers: The social value of avoiding health decline: An experimental approach. BMC Medical Informatics and Decision Making 2002;2(3).

(24) Scott A. Identifying and analysing dominant preference in discrete choice experiments: an application in health care. Journal of Economic Psychology 2002;23:383-398.

(25) Lancsar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Econ 2006;15:797-811.

(26) Dolan P, Edlin R, Tsuchiya A, Armitage C, Hukin A, Brazier J, et al. The relative societal value of health gains to different beneficiaries. 2008;RM03/JH11.

(27) Mason H, Baker R, Donaldson C. Understanding public preferences for prioritizing health care interventions in England: does the type of health gain matter? Journal of Health Services Research & Policy 2011;16:81-89.