CHOOSING A HIGH-QUALITY HOSPITAL

The role of nudges, scorecard design and information

Tammy Boyce
Anna Dixon
Barbara Fasolo
Elena Reutskaja

The King's Fund>

Contents

List of figures and tables

About the authors

Tammy Boyce is a freelance researcher and writer. At the time of writing this report, she was the research fellow in public health at The King's Fund. She also worked for the Marmot Review of Health Inequalities and at Cardiff University's School of Journalism. She is the author of *Health, risk and news: The MMR vaccine and the media* and is co-editor of *Climate Change and the Media.*

Anna Dixon is Director of Policy at The King's Fund. She has conducted research and published widely on health care funding and policy. She has given lectures on a range of topics including UK health system reform and patient choice. She was previously a lecturer in European Health Policy at the London School of Economics and was awarded the Commonwealth Fund Harkness Fellowship in Health Care Policy in 2005–6.

Anna has also worked in the Strategy Unit at the Department of Health, where she focused on a range of issues including choice, global health and public health.

Barbara Fasolo is Lecturer in Decision Sciences at the London School of Economics and Political Science. She is affiliated with LSE Health (Risk Research Unit) and the LSE Decision Capability Unit. Since 2009 she has a part-time secondment at the European Medicines Agency to study and aid drug regulators' quantification of benefits, risks and their balance. Her area of expertise is behavioural science and experimental decision-making and she has studied decisions across a range of domains. She publishes and has given talks in areas of behavioural decision science as applied to policy-making, regulatory decision-making, decision analysis, financial and managerial decision-making.

Elena Reutskaja is assistant professor in the Marketing Department at IESE Business School. Professor Reutskaja's research lies at the intersection of behavioural decision-making and marketing. She studies how people make decisions in their daily lives from choosing a snack to selecting a hospital for a treatment, focusing on the problem of too much information and choice overload. Professor Reutskaja is actively involved in a number of international research projects both for the private and public sectors. She was recently a visiting scholar at the California Institute of Technology, where she was engaged in brain-imaging and eye-tracking studies with applications in economics, business and policy-making.

Acknowledgements

Executive summary

The coalition government is committed to increasing choice and to publishing more information about the quality of care provided by health care organisations. In future, most of this information will be published electronically on the internet. The 'information revolution' aims not only to increase transparency but also to inform patient choice, and this, in turn, is expected to drive improvements in the quality of care.

This research examines some of the ways that organisations can help patients to make better use of clinical quality information when deciding which hospital to attend. It is based on a year-long study using focus group discussions and a series of online experiments.

Rational decision theory argues that if individuals are adequately informed, they will make choices that maximise their own interests. However, in practice, when faced with complex decisions, people tend to abandon logic and rely on intuition to guide their decisions. There is increasing interest in various policy circles in the idea that you can use 'nudges' to help people make better choices. In this research, we designed an interactive website in which people were presented with information about different hospitals. We experimented with several different nudges with the aim of helping people to pay more attention to information about the clinical quality of services, and to use this information to choose a hospital.

We aimed to answer the following questions:

- What information do patients use when choosing a hospital?

- What is important to patients when choosing a hospital?

- How does the design of information influence the choices that patients make, and in particular, how do nudges affect people's decisions?

- Do people make better choices as they become more practised in making decisions (ie, do they learn to make better choices)?

Type of information

People can be easily overwhelmed by the quantity of information available, and spend only a comparatively short amount of time (a few minutes in many cases) consulting this sort of information online. Websites provide opportunities to filter information so that people are only presented with information on those aspects of care that they indicate are most important to them.

We found that while older people (those aged 51 and over) and those with lower levels of education (who had only attained secondary education) were keen to have summative measures, they were also confused by their meaning. Further testing is needed to establish how to combine summative and disaggregated indicators effectively for different people.

Participants in our focus groups were suspicious of terms such as 'missing data' or 'data not available'. Information providers need to improve the presentation of missing data, explaining what 'data not available' means and why it is not available.

Focus group participants also made it clear that they wanted information about the individual consultant or doctor who would be treating them, information that was relevant to their condition or treatment, and which helped them to understand the risks they faced by attending a particular hospital. This suggests the need for information at individual clinical level – information not currently available, except in relation to cardiothoracic surgeons.

Preferences

We found that people do not have stable preferences about what is important to them when choosing a hospital. This suggests there is an opportunity for information providers to influence what patients pay attention to by making some aspects of hospital care, such as safety or quality indicators, more salient.

Coaching people about their preferences before making a decision made people use information more systematically, and they were more likely to compare hospitals on the indicators they thought were important rather than trying to take in information about each hospital.

Decision aids or scorecard designs that elicit preferences and values may help patients to identify what is important – that is, to shape preferences. However, the ability to personalise information so that the support offered can help people to choose 'the best for me' is still some way off.

Presentation

The way information is presented can make a difference to how it is used. We made a number of improvements to the design of the scorecard used in the online experiment, which our focus groups and previous research suggested would help.

- We used clear, easily understood labels for indicators that did not require people to 'drill down' for further information.

- We used intuitive symbols and labels consistently (ticks rather than traffic lights or colours).

- We applied evaluative labels – eg, 'poor', 'good', 'excellent'.

These design features should be adopted by information providers who are presenting comparative quality information for use by the public and patients, particularly for the purpose of making choices between providers.

Information about the quality of health care, whether on organisations or individuals, often involves the use of numerical information, and in some cases, such as the Hospital Standardised Mortality Rate, this involves quite complex statistical information. Our research confirms that only people with high levels of numeracy are able to process this information in order to make a choice.

Ordering

It matters what information you put first. We designed the online scorecard so that the quality measures were first, and these were viewed most frequently, according to the heatmaps.

We also thought that by sorting the options (in this case hospitals) by quality, we might also make it more likely that patients would choose the highest quality hospital. Although sorting hospitals by quality might appear to be a good idea, it actually resulted in people making worse choices. In fact, those who had the options presented by distance did better. Clearly, sorting can have perverse effects, and needs to be carefully tested.

Use of nudges

We were interested in whether different nudges can improve the choices people make, how satisfied people are with the choice they made, and whether nudges can help people to make more informed choices in future. No nudge performed well on all these criteria.

We used three nudges (in addition to sorting the hospitals by quality and distance): a pre-select nudge (where the highest quality hospital was pre-selected), a preference nudge (where participants were asked to rank the indicators before they were presented with the scorecard) and a feedback nudge (where participants were asked to reconsider their choice if they had not chosen the highest quality hospital).

The pre-select nudge performed worse in terms of choosing the highest quality hospital, but people who had this scorecard did best in terms of making a choice when presented with a similar task a few weeks later, when nudges were not used. The participants who had the pre-select nudge were also quite dissatisfied with their choice, compared with those who had a very simple scorecard (sorted by distance or quality).

Giving people feedback on their selection and asking them to reconsider can increase the proportion that will choose a high-quality hospital, but it does not necessarily help people to make better choices in future, without nudges present.

Making people think about what was important to them in a hospital, and that the quality of hospitals varied, meant that fewer were very satisfied with their choice, and these people did worst when presented with a similar task a few weeks later.

The results also show that nudges have different impacts on different people – the feedback nudge made it more likely that elderly participants would choose the highest quality hospital, but in general, younger participants, particularly those with higher education, appear to have been helped more by the nudges. It appears from the findings that younger participants benefited from information about differences in quality of care and being made to think about what was important to them before making the choice.

Nudges require the information provider to decide on what they are seeking to nudge people towards, and therefore, a normative decision is inherent. However, the 'best' hospital on average may not be the 'best' choice for any particular individual.

Practice

Choosing a hospital for a surgical procedure is not a task that people face frequently. The results of this research suggest that repeating the exercise meant that people were more likely to choose the highest quality hospital, even when asked to make a choice of hospital again a few weeks later. This applied regardless of age and education. It is likely

that younger people are less familiar with health care, so giving them the opportunity to 'practise' making a choice increases their awareness of the factors that might be important. This is consistent with the impact of giving people information before they were asked to make a choice, which appeared to help young people most. For older people, while they are more familiar with hospitals, they are perhaps less familiar with the task – that is, using the internet to compare products or services. So giving them the opportunity to practise helped them make a better choice.

Conclusions

This research reinforces the importance of paying attention to what information is presented and how, and cautions against a mantra that 'more information is always better'.

People find it difficult to make trade-offs between quality, safety, patient experience and location. Our findings suggest that the government should be cautious about the ability of patients (apart from those who are highly numerate) to make these complex decisions without some decision support.

Making people more aware that the quality of hospitals differs, and giving them opportunities to practise making a choice, appears to help people make better decisions. More research is needed to evaluate the effects of different nudges on patient decision-making.

It appears that exposing people to some of the differences in quality between hospitals and forcing them to consider these difficult trade-offs may increase their dissatisfaction with the choice they make. Patients may benefit from information that reassures them that hospitals meet a minimum set of required standards.

There needs to be an evidence-based approach to the public reporting of comparative performance information in future. Simply allowing all the information currently held about the quality of care to be put in the public domain will not result in people making informed choices.

Introduction

Imagine you need to buy a car. For the same price, you have a choice of two:

Car A:
- high quality
- high safety
- car dealer is far: need to travel to the outskirts of town

Car B:
- low quality
- medium safety
- car dealer is close: down the road from where you live.

Most of us do not know how a car works, but we use other skills and information to weigh up the pros and cons and make a final decision about which car to buy. In the choice between A and B above, if information on the car's quality and safety were given, most would pay attention to that rather than to the proximity of the car dealer selling it. Paradoxically, when people have to make more complex decisions – such as which hospital to be treated in – they find it much more difficult to identify what really matters. Patients might therefore find it more difficult to weigh up issues around the quality and safety of hospital services with how far to travel. As a result patients might rely on factors such as ease of access (distance from their home) at the expense of other important factors like clinical quality.

> *Making trade-offs in order to integrate conflicting dimensions into an overall choice is difficult… Consumers use simple heuristics, eg, give more weight to precise variables such as costs [or location] and less weight to fuzzy variables such as quality.*
>
> (Hibbard *et al* 1997)

The theory of rationality argues that if individuals are adequately informed, they will make choices that maximise their interests. However, when faced with complex decisions, and lacking expertise, people tend to abandon logic, rationality and information, and rely on intuition to guide their decision. In situations that are unfamiliar and that offer little opportunity for feedback, decisions based on intuition alone lead to a series of common and predictable errors.

Thaler and Sunstein (2008) have argued that these errors can be exploited, and one can help people to make better choices by designing 'nudges' that make it easier for people to do the 'right' thing. Policy-makers have embraced the concept of nudging with enthusiasm, using nudges in a range of contexts (Dolan *et al* 2010), especially those that require no reflection, like defaults (Halpern 2009).

This research addresses issues around informed choice in the context of health care. It examines how patients make decisions about which hospital to choose for treatment, and how they can make better use of information on clinical quality when making their decisions.

In recent years, there has been rapid growth in the amount of information available to patients on the quality of treatment and care, collected within the National Health Service (NHS) both routinely and as part of clinical audits (Raleigh and Foot 2010). Initially, these data were published for the primary purpose of performance management, benchmarking, and so-called 'naming and shaming' by regulatory organisations such as the Healthcare Commission. However, data have also been published by commercial organisations like Dr Foster and the government-funded website NHS Choices, for the purposes of informing patient choice.

High Quality Care For All (Department of Health 2008), published under the Labour government, included a commitment to increase the amount of information on quality and safety that is collected and publicly available. It highlighted the importance of making such information available to the public and patients, as well as it being used by clinicians to understand how they are performing relative to others, as a driver of quality improvement. Subsequently, legislation was introduced to require all providers of NHS-funded care to publish annual quality accounts. The first set of quality accounts was published in July 2010 and is the subject of a separate analysis by The King's Fund (Foot *et al* forthcoming).

The coalition government has emphasised the importance of giving patients more choice, including choice of GP. Yet, as others have noted, 'choice is only meaningful if patients know about the options available to them and can compare these effectively' (Ellins and McIver 2009). It is, therefore, not surprising that there is also a strong commitment to publishing more information. Public reporting of hospital performance data is widely supported as a means of improving quality in the NHS. Andrew Lansley, the Secretary of State for Health, in his first major speech on the NHS, said:

> *For patients, they need to know who is providing quality, safe, effective, accessible services. Information will drive higher standards. It's not just about choice, although patients value choice, even if the choice they make continues to be their local practice and their local hospital. But information and choice hold people to account. So our vision will be of an information revolution across the NHS… Putting the information out there – accessible to everyone – is a catalyst. It drives comparison and performance.*
>
> (Lansley 2010)

In its programme for government, the coalition committed to 'publish detailed data about the performance of health care providers online, so everyone will know who is providing a good service and who is falling behind', as well as enabling patients to 'rate hospitals and doctors according to the quality of care they received' (HM Government 2010). The recently published White Paper, *Equity and Excellence: Liberating the NHS* (Department of Health 2010a), and accompanying consultation on the outcomes framework, have set out in more detail the government's ambitions in this area. The policy focus on patient-centred services, accountability to patients for outcomes, and increased transparency, suggests that recent developments within the NHS – to collect and publish information on the performance and quality of organisations – will continue, and potentially accelerate. However, the White Paper pulled back somewhat from the idea of rating and comparing individual doctors, referring instead to clinical departments and teams.

Increasingly, information about quality, safety and performance is being published on the internet. NHS Choices is the main government-funded website that provides information to patients. It aims to 'empower the public to *make informed choices* [our emphasis] about their own health, including when and where they receive treatment' and to provide '*easily accessible* [our emphasis] information about scorecards, treatments and hospitals' to help patients make informed decisions (NHS Choices 2010). There is considerable research

on how consumers use online product comparison sites, as well as how patients use and understand comparative information on hospital quality. Our research aims to build on this knowledge and explore how the presentation of comparative quality information online affects how people make decisions, and how patients make choices about where to go for treatment.

Concerns have been raised about whether the emphasis on patient choice and web-based information disadvantages certain groups, who may find it more difficult to access and interpret complex information. Previous research suggests that different types of information are needed to meet the needs of different groups – for example, the young, older people, and those with different levels of education (Ellins and McIver 2009). So we also examined how different groups of people use information – elderly people, those with lower levels of education, and those less familiar with the internet – as well as whether the presentation of comparative hospital indicators affects their decisions, and how.

The research aims to inform future policy on the publication of quality information for patient choice, and to provide some practical suggestions for how this information should be presented to help patients make informed choices on the basis of the clinical quality of providers. While patients may make a choice that is consistent with their personal values and preferences (for example, a local hospital), this may not be the best choice from a system perspective (for example, data may suggest this hospital has a poor clinical or safety record). In order for the market in health care to operate effectively – that is, where competition drives quality improvement – patients need to choose on the basis of differences in clinical quality. We are concerned with how information can best be presented to make it more likely that patients will choose high-quality providers.

This report aims to answer the following key questions:

- What information do patients use when choosing a hospital, and how do they use it?

- What do people say influenced their decisions when choosing a hospital?

- What choices do people make, and who makes the 'best' choices?

- How does the way information is designed influence the choices that patients make, and, in particular, how do nudges affect people's decisions?

- How satisfied do people feel about their decisions?

- How aware are people of differences in quality between hospitals, and what trade-offs do they make in their decisions?

- Can scorecards be improved to teach people to be informed consumers/make better choices?

We carried out a year-long investigation, drawing on focus group discussions and a series of online experiments. The next section outlines the current research on decision-making, particularly with regard to comparative hospital indicators. Section 2 describes our methodology and outlines the qualitative and quantitative methods used. Section 3 reports our findings from the focus groups, while Section 4 details findings from the online experiments. The final section discusses the implications for policy-makers and information providers.

1 Background

Policy context

Within the NHS, patients have traditionally had limited opportunities to make choices, at least compared with other health systems such as in France and Germany, and patients with health insurance in the United States. In England, GPs have been free to refer patients to any NHS hospital since the NHS was created (apart from a brief period when they were limited to those hospitals the health authority had a contract with).

It is only in recent years that policy has focused on supporting patients to make choices about where they see a specialist and where they receive treatment, and ensuring that information is available to enable them to make the best choice. Information on the performance of hospitals has been published for a number of years, but initially, this was not specifically for the purposes of patient choice. Most information was produced to support performance management (and later, regulation) in the NHS; it was published as part of a strategy to make performance more transparent and to drive improvement through 'naming and shaming'. These two developments – strengthening patient choice and greater transparency of performance information – have come together most recently in an active policy to promote informed choice of hospital. The right to an informed choice was enshrined in the NHS Constitution in 2009 (Department of Health 2010b).

Patient choice was piloted in a limited way in London (from October 2002 to April 2004) and for cardiac surgery patients across England (from July 2002 to November 2003). The focus at this time was on patients who would otherwise wait for treatment; they were given the choice of being treated more quickly at an alternative provider with lower waiting times and/or spare capacity. Patient choice for those facing a wait of six months or longer was rolled out across England from 2004; however, as waiting times were falling rapidly, the number of patients eligible for choice diminished. From January 2006, patients were given a choice of at least four providers at the point of referral for a specialist outpatient consultation. Since April 2008, choice was further extended to include any NHS or independent sector provider (registered with the Care Quality Commission (CQC)) willing to supply care at the fixed national price (tariff).

Patient choice, together with other policy changes, was intended to create competition between hospitals for patients, with the expectation that this would lead to improvements in efficiency, equity, quality and responsiveness. Whereas in the early phase of the policy, the key criterion on which patients were expected to choose was waiting times, policy-makers expected that patients making a choice at the point of referral would pay more attention to clinical quality. But research shows that a patient's personal experience, and that of their family and friends, remains important when making a choice (Dixon *et al* 2010).

Recent developments in online information mean that feedback from other patients about their experience of a hospital is now easily accessible. However, there remains policy interest in encouraging patients to pay more attention to clinical quality (as opposed to patient experience) when making a choice. The theory underpinning patient choice policy is that patients will choose to be referred to higher quality hospitals; providers

delivering lower quality will lose patients – and the money that follows them – under activity-based payment (Payment by Results). The theory is that quality will improve as hospitals seek to attract (or retain) patients (Hirschman 1970; Hibbard *et al* 1997; Faber *et al* 2009; Lindenauer 2009).

Information on the performance of NHS trusts was first collected as part of the performance assessment framework in 1999 (Raleigh and Foot 2010). The Department of Health published the first star ratings of trusts based on similar data in 2001. This was controversial, but has continued, albeit under the guise of the Annual Health Check, published first by the Commission for Health Improvement (CHI) and then the Healthcare Commission. The CQC will no longer publish an Annual Health Check, but instead produces quality and risk profiles on each provider. These will become available to providers from autumn 2010, and some of this information will be published from January 2011. The CQC increasingly collects and publishes data on the quality of care delivered by private sector providers. It has implemented common standards of care across public and private providers, so that in future, there should be more comparable data.

There is less data available to compare the performance of individual doctors. Data on the outcomes of cardiothoracic surgeons have been published since 2004 (Bridgewater and Keogh 2008) (*see* Figures 1, below, and 2, overleaf). They remain an exception among clinical specialties, which have generally not taken the initiative to publish outcomes.

Figure 1 Screenshot of website showing information on survival after heart bypass surgery
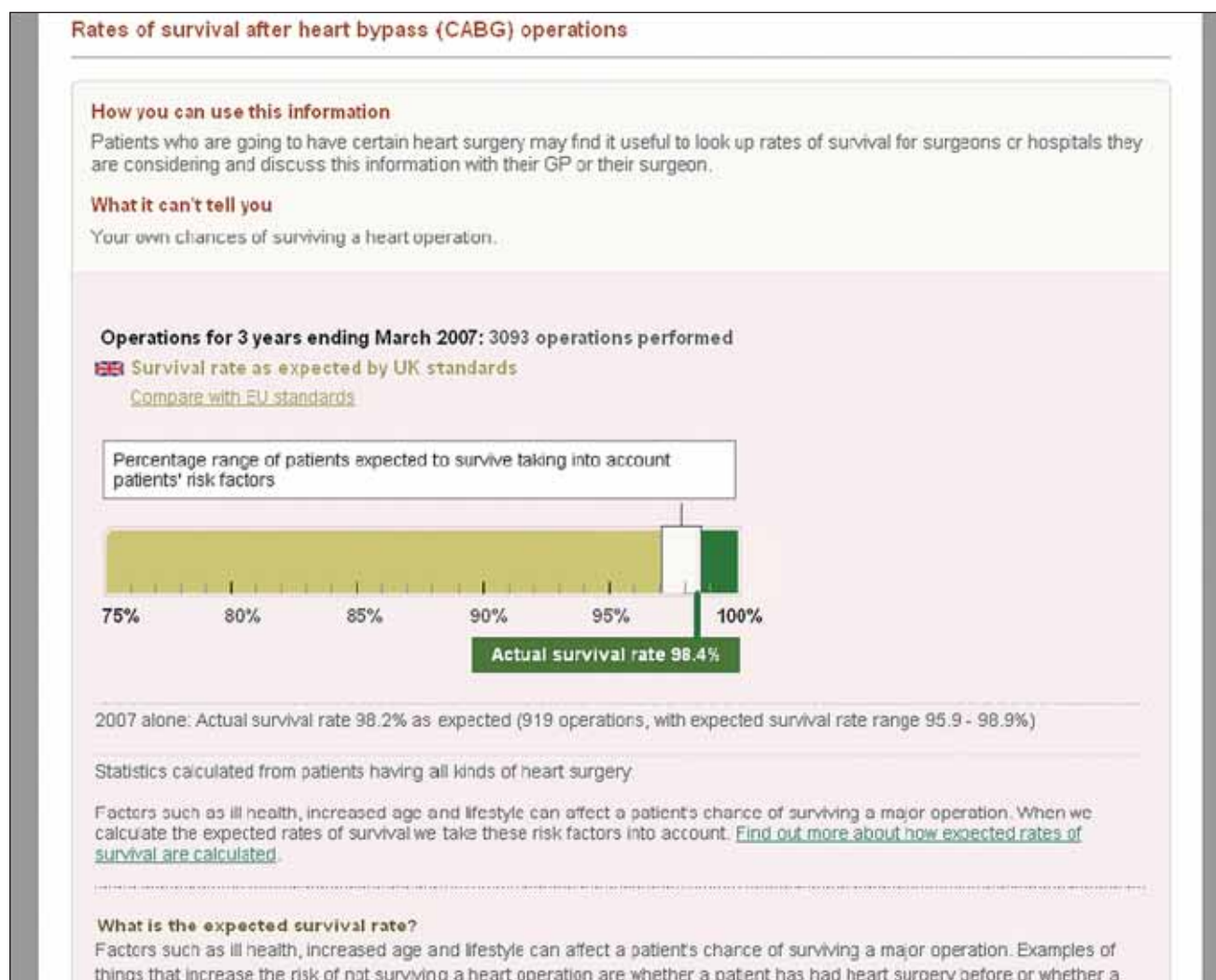
Figure 2 Screenshot of website showing number and type of cardiac surgery procedures undertaken by individual surgeons

The NHS Choices website has roughly 80,000 pages of information (NHS Choices 2010), and recorded 100 million visits over the past 12 months (NHS Choices 2009). It provides comparative hospital performance indicators, as well as tools to help patients manage their own health and well-being. The 'find and choose' services section allows people to search by postcode. The results are presented either in list form or sorted by distance ('detailed comparison') (*see* Figure 3 opposite). Patients can select a particular scorecard or procedure for which they are being referred, enabling them to see information specifically relevant to them.

NHS Choices introduced a facility for patients to provide their own feedback and views on hospitals in 2007. It has since expanded the facility to include GPs and, by the end of 2010, patients will be able to leave feedback on a wide range of services, including dentistry and maternity care. Other websites such as iWantGreatCare and patientopinion.org.uk also allow patients to share and compare their views on health care providers and individual doctors.

Figure 3   Screenshot of NHS Choices website showing 'detailed comparison' of five hospitals sorted by distance

We found **87 hospitals** within 50 miles ▾ of w1g 0an   for **Hip replacement**

▸ **Refine your results**

Update results ›   › Start a new search

Results: 1 - 5 of 87   Order by: closest shown first ▾   Page view: Simple list | Detailed comparison | Map

The Heart Hospital	**University College Hospital**	**St Mary's Hospital (HQ)**	**St Thomas' Hospital**	**Guy's Hospital**
(providing services for University College London Hospitals NHS Foundation Trust)	(providing services for University College London Hospitals NHS Foundation Trust)	(providing services for Imperial College Healthcare NHS Trust)	(providing services for Guy's and St Thomas' NHS Foundation Trust)	(providing services for Guy's and St Thomas' NHS Foundation Trust)
16-18 Westmoreland Street London Greater London W1G 8PH | 020 7573 8888 | | 0.26 miles from w1g 0an	235 Euston Road London Greater London NW1 2BU | 0845 155 5000 | | 0.67 miles from w1g 0an	Praed Street London Greater London W2 1NY | 020 3312 6666 | | 1.24 miles from w1g 0an	Westminster Bridge Road London Greater London SE1 7EH | 020 7188 7188 | | 1.76 miles from w1g 0an	Great Maze Pond London Greater London SE1 9RT | 020 7188 7188 | | 2.61 miles from w1g 0an
Add to shortlist	Add to shortlist	Add to shortlist	Add to shortlist	Add to shortlist

▸ **Number of operations performed**

▸ **Waiting times**

▸ **Average time spent in hospital**

▾ **Patient outcomes for this operation at these hospitals**

The Heart Hospital	*University College Hospital*	*St Mary's Hospital (HQ)*	*St Thomas' Hospital*	*Guy's Hospital*
As expected The rate of unplanned	**As expected** The rate of unplanned	**As expected** The rate of unplanned	**As expected** The rate of unplanned	**As expected** The rate of unplanned

Before we describe the research design and methodology (in Section 2), we summarise the theoretical and empirical literature on how patients use comparative performance information. The majority of research studies on the effects of comparative indicators are from the United States and assume that 'there is a competitive market with multiple providers or health plans and free choice' (Faber *et al* 2009, p 6). The effectiveness of comparative indicators in state-funded health care systems such as the NHS (where cost is not a direct concern for patients) is less well understood. In addition, most empirical studies have examined the impact of paper-based reports, rather than online information.

Theories of decision-making

Traditional economists and policy-makers often assume that patients make optimal decisions (that is, those that maximise their utility) when they are adequately informed. However, making decisions concerns more than having access to good information. According to dual-process theories (Chaiken and Trope 1999), people typically reason in one of two ways:

■ automatic decision-making: associative, fast, uncontrolled, unconscious, uses shortcuts (intuitive heuristics) (system I). Heuristics are mental shortcuts used to help people solve problems and make decisions quickly and efficiently. They are helpful in making quick decisions, but carry the risk that potentially relevant information is ignored, and therefore poor decisions or 'decision errors' are made (Tversky and Kahneman 1974). For example, people may make decisions based on cost rather than quality because such information is precise, direct, and easy to compare.

■ reflective decision-making: rule-based, conscious, deliberative, controlled, effortful, slow (system II) (Stanovich and West 2000).

System I over-rides system II whenever decisions are complex, decision-makers are not experts, and there is time pressure.

The design of scorecards (referred to as report cards in the United States) commonly assumes that people make decisions deliberately. Yet, when people are overwhelmed with information, they cannot easily discover what is most important to them, and they lack the skill and knowledge to make trade-offs, so they base their decision on intuition and fast heuristics. There are also examples from consumer research where consumers who encounter complex information do realise the need to compare information and want to make tough trade-offs, and try, as well as they can, to make such comparisons both online (Fasolo *et al* 2005) and offline (Bettman *et al* 1993) though they find such comparisons challenging.

People find it difficult to make decisions about health care: the information is complex, the decisions are important, and carry high risks. Therefore, they need information that is easy to understand in order to make informed choices (*The Lancet* 2005; Ellins and McIver 2009). Online scorecards and similar comparative information on hospital performance require people to compare a large number of indicators and make complex trade-offs; these skills are cognitively difficult and emotionally draining, even for young and highly numerate people (Bettman *et al* 1993; Luce 1998). The skills needed to make decisions using comparative quality information include the ability to:

■ correctly interpret data

■ identify important factors to integrate into a decision

■ weight factors in ways that match one's individual needs and values

■ make trade-offs

■ bring all factors together into a choice (Hibbard *et al* 2001).

People also need to be able to define their individual preferences. Previous research has shown that the availability of an ideal alternative, or the ability to understand what characteristics the ideal alternative should have, can simplify complex decisions (Chernev 2003). However, compared with other consumer choices – where consumers are able to set clear preferences before they make a choice – many patients are not experts in making decisions about hospital choices, and fail to understand what exactly they want, both before receiving information, and while they are being provided with information. Difficulty in identifying one's own preferences makes decisions about which hospital to choose even more complex.

When people are faced with an overwhelming amount of information, or information they are unfamiliar with, they use various methods to help them make choices:

Consumers facing complex, high-stakes choices are prone to predictable errors. They are likely to lack the skill and time to make choices based on a careful assessment of the relative costs and quality of competing health plans, tending instead to choose on the basis of anecdotal information, such as their friends' experiences.

(Frank and Zeckhauser 2009)

Increasingly, websites such as NHS Choices mix 'soft' data such as anecdotal information or patient feedback with 'hard' data. Because personal experience is compelling, there is a risk that this carries more weight than objective data. Indeed, there is some evidence from medical decision-making that the use of patient testimonials, even where these are 'representative' (for example, if there is a one in four chance of a successful outcome, patients are shown one negative testimonial and three positive ones), over-rides any statistical information (Ubel *et al* 2001; Fagerlin *et al* 2005). Based on this literature, the message seems to be that if you want people to look at 'objective' medical information, it is best not to accompany this with testimonials; otherwise, patients will use inaccurate heuristics to make their decision. We did not, therefore, include anecdotal information; the data on patient experience was based on routinely available surveys and also presented as 'statistics'.

The amount of information provided also significantly affects decision-making. People often claim they want more information when they evaluate options, yet research consistently finds that 'less is more' – that is, people make better evaluations and decisions when they are presented with less information about their options (Hibbard *et al* 2001; Hibbard and Peters 2003; Peters *et al* 2007a). However, some people may want more information, even if they feel overwhelmed by it – because having more choice is associated with greater decision freedom, autonomy, self-control and intrinsic motivation (Zuckerman *et al* 1978; Walton and Berkowitz 1979; Ryan and Deci 2000). Therefore, there is a tension between the amount of information consumers say they want and the amount they actually use to make their decision (Damman *et al* 2009).

On average, people are able to process approximately seven pieces of information (plus or minus two) at a time (Miller 1956). Presenting too many choices can lead to 'options overload', dissatisfaction, and even choice deferral (Schwartz 2004). Consumer comparison websites are aware of this, and while there is a large number of products, consumers are only 'allowed' to compare between three and seven items (Edwards and Fasolo 2001). By allowing only a limited choice, people then experience 'freedom of choice' without feeling overloaded (Gourville and Soman 2005). People are generally more satisfied when they are presented with fewer rather than more choices (Iyengar and Lepper 2000; Schwartz 2000).

Importantly, in most consumer situations, and especially where people are unfamiliar with the information presented, they are unlikely to have well-defined or fixed preferences about what is important to them or what their priorities should be. As a result, preferences or opinions are constructed 'on the spot' (Lichtenstein and Slovic 2006), and powerfully shaped by the context. However, getting people to think about their preferences and articulate them before choosing (for instance, forcing people to think about what constitutes an 'excellent quality' hospital) helps to simplify choice (Chernev 2003).

Understanding how patients make choices when presented with comparative information on hospitals is important to ensure that patients are supported to make high-quality decisions and are not overwhelmed by information.

Previous research

Awareness and use of scorecards

People are increasingly aware of their right to make choices about their health care – half are aware of the right to choose where they are treated for non-urgent care (Appleby and Phillips 2009). Awareness of patients' right to choose is increasing each year, but remains low. In 2004, 9 per cent of patients knew they had a right to choose which hospital they were treated in, increasing to 19 per cent in 2007 (Appleby and Phillips 2009). Older people and those with higher levels of education are more likely to be aware of their right to choose (Dixon *et al* 2010).

Data from the Department of Health monitoring survey shows that about half of patients are aware of their right to choose (this has risen from 29 per cent in May/June 2006) (Department of Health 2009). Less than 20 per cent of respondents who were referred to an NHS provider were aware that they could choose to have NHS treatment in a private sector hospital, compared with 41 per cent among those who were referred to an independent sector treatment centre (ISTC) (Dixon *et al* 2010).

Awareness of the NHS Choices website remains stubbornly low (2 per cent spontaneous awareness and 21 per cent brand recognition in an omnibus survey (N=1,750)), despite efforts to promote and publicise it. In a survey of patients recently referred for an outpatient appointment, only 4 per cent used NHS Choices to help them choose. Seven per cent said they had heard about the performance of local hospitals from published performance reports (Robertson and Dixon 2009). Despite low levels of reported use, NHS Choices received more than 100 million visits in 2009 (Bob Gann, personal communication, August 2010).

There is considerable demand for accessing health information online. People report that health websites are a useful source of information that is not routinely provided by health professionals (Nicholas *et al* 2004; Coulter and Ellins 2006). Yet only 6 per cent of the patients in the latest Department of Health survey reported that they had used NHS Choices when asked this question: 'Which, if any, of the following sources of information did you use to choose the hospital?'; and just 1 per cent turn to other comparative health care websites (Department of Health 2009). Patients are more likely to rely on their GP or the experiences of their friends and family to help them choose a hospital (Department of Health 2009; Dixon *et al* 2010). Levels of awareness and use of scorecards are also low in the United States. In one survey, only 7 per cent of participants said they had seen and used information comparing the quality of hospitals to make health care decisions in the prior year (Shekelle *et al* 2008; Lindenaur 2009).

Research has sought to understand why comparative indicators are not used more often, particularly as people become more dependent on the internet to help them find information and make choices.

Evidence suggests that people frequently find scorecards too complex and detailed, with too much information to process or use effectively (Marshall *et al* 2000, 2006; Wensing and Grol 2005). Users also find the route to information too challenging and do not understand the information presented (Dr Foster 2007). More specific research examining why NHS Choices is poorly used has found that people are suspicious about the Department of Health (distrusting government statistics) and the lack of detail 'about specific local services rather than generalised comparative information' (Ellins and McIver 2009).

Reviews carried out by Marshall and colleagues (2000, 2006) found a number of reasons why patients did not use information about the quality of health care. These included:

- difficulty in understanding the information

- disinterest in the nature of the information available

- lack of trust in the data

- problems with timely access to the information

- consumers rating anecdotal evidence from family and friends more highly than empirical evidence.

People appear to be unaware of where to find information to help them compare the quality of service available and to make health care decisions. For those who find the relevant websites, the information is difficult to understand.

What information do patients want to help them choose?

Much of the research on the type of information patients want when choosing health care is based on experience in the United States, which there is a longer history of using scorecards than in the United Kingdom, both on paper and online. Research on the US market finds that patients primarily base their choice of hospital on quality factors and clinical reputation, doctor's recommendation, and location. However, patients are also concerned with non-clinical indicators, including comfort of the rooms, or registration procedures (Grote *et al* 2007). In one survey, the majority of respondents (59 per cent) stated they would continue to choose a hospital that is familiar to them rather than a higher rated facility (Lindenauer 2009).

Research from the Netherlands reflects these findings, concluding that despite the vast amounts of information about quality found in online scorecards, many patients base their health care decisions on their own experiences or those of their family and friends (Damman *et al* 2009). Integrating hard performance data with softer sources of information such as patient experience has been identified as an important influence on patients' decisions (Marshall *et al* 2006; Damman *et al* 2009; Moser *et al* 2010). Other consumer research has found that when both types of information are available, consumers often end up comparing and choosing from the subset of options which are most highly recommended (to avoid those that had poor reviews), or try to find if there is one product that beats the most highly recommended. If the products are high risk (as hospitals would be), people rely more on recommendations as time pressure increases (Rosen and Olshavsky 1985, 1987a, 1987b).

In England, a number of studies have analysed what factors patients consider to be important when choosing a hospital. Discrete choice experiments have found that patients wanted information on mortality rates, readmission rates, and hospital-acquired infection rates, as well as waiting times and GP opinion, but they were still more likely to consistently choose their local hospital (Burge *et al* 2005, 2006). Dixon *et al* (2010) carried out a discrete choice experiment, asking patients to weigh up the performance of available hospitals. They found that distance was an important factor influencing choice of hospital, with 61 per cent preferring to be treated in their local hospital. Even though people say that they are interested in quality indicators, in the United Kingdom, they still tend to choose hospitals based on shortest distance from their home and shortest waiting times (Burge *et al* 2005; Lim and Edlin 2009). In England, there are signs that this attitude is changing, as location is becoming less important, and quality more important, when choosing hospitals (Coulter *et al* 2005; Audit Commission and Healthcare Commission 2008; Dixon *et al* 2010).

Other factors influence the type of choices patients make. When patients are offered a choice, most choose their local provider; but patients aged 51–65, those living outside of cities and large towns, those educated to degree level, and those with a bad past experience of the local hospital are more likely to choose care from a non-local provider (Dixon *et al* 2010). When the analysis was repeated for those not offered a choice, the only variable that affected whether patients travelled to a non-local provider was their past experience of the local hospital. Using data from hypothetical choices, studies have found that people with low levels of education were less influenced by increases in clinical quality compared with those with higher levels of education (Burge *et al* 2005, 2006). Having poor health and being reliant on bus transport was associated with greater loyalty to their local hospital; whereas those with higher education, internet access and a poor perception of the local hospital were more likely to choose a hospital further away (Dixon *et al* 2010).

The key is to find a balance between what people want and can cope with (bearing in mind that different demographic groups will have different needs), and what information best helps people to make informed decisions. This may involve prompting people to be interested in certain factors.

Strategies to improve presentation of scorecards

Information on scorecards should be presented in a form that makes it easy to process. This means using strategies that lower the cognitive effort required, helping people imagine what the actual experience of choice might be like, and finally highlighting the meaning of the information (Hibbard and Peters 2003). Processability is determined by:

- limited cognitive resources (attention, memory, time) (Miller 1956)

- limited processing speed

- conflicting goals when making trade-offs (maximise accuracy and transparency, and minimise effort and emotional strain) (Payne *et al* 1993; Luce *et al* 1999; Fasolo *et al* 2005).

A systematic review of public reporting found that reports increased people's knowledge, altered attitudes, and changed choices made in laboratory scorecards; but in real-world experiments, these reports had no effect on knowledge, attitude or behaviour (Faber *et al* 2009, p 5).

Decisions are heavily influenced by how they are framed and contextual factors such as the ordering of information (Damman *et al* 2009). Therefore, the way information is presented, both the order and the measurement labels used, influences how that information is received (Schkade and Kleinmuntz 1994; Lohse 1997; Thaler and Sunstein 2008). For example, information listed first is generally considered first and given more weight than information given further down a list, especially if the list is long (Lohse 1997).

People find it difficult to decipher meanings when different formats are used to compare hospitals and hospital treatment (Vaiana and McGlynn 2002; Burgess *et al* 2005). Easy-to-read visual ratings such as stars, bars, letter grades or numbers are considered the easiest to understand and use (Hibbard *et al* 2001; Hibbard *et al* 2002; Uhrig *et al* 2006; Peters *et al* 2007a).

The way that patients use and interpret information is also affected by demographic characteristics. Certain groups, such as elderly people or those with low numeracy, benefit most from the use of visual representations (Hibbard *et al* 2001; Peters *et al* 2007a). Those

with low levels of reading skills and numeracy may require additional support to make informed choices about comparative hospital care. A 2003 *Skills for Life* survey found that one in six working-age adults (16–65 years), equivalent to 5.2 million adults in England, have reading skills at or below the level expected of an 11-year-old, rising to one in two (15 million adults) for numeracy (Department for Education and Skills 2003).

Low levels of literacy and numeracy therefore impede the use of comparative health information and may potentially exacerbate health inequalities. The barriers are highest for those on low wages, with lower education, and from minority ethnic backgrounds (Faber *et al* 2009). Poor numeracy skills were the largest factor inhibiting patients' ability to use comparative performance information (Hibbard *et al* 2007). Those with lower levels of education and those with English as a second language already have difficulty understanding health care information, and this can further exacerbate their effective use of comparative performance information (Fagerlin *et al* 2007; Ellins and McIver 2009).

Numeracy is not the only factor potentially contributing to inequalities in the use of online health care comparison websites. Digital exclusion may be an additional impediment. Although the Department of Health produced a booklet on choosing a hospital, containing information about local options, which GPs are required to give to patients, few patients report receiving this booklet (Department of Health 2009). People do not use the internet as their main source of comparative information on hospital performance. Recent estimates suggest that as many as 10.2 million adults (21 per cent of the UK population) have never accessed the internet, and 30 per cent of households are without an internet connection (PricewaterhouseCoopers 2009). Among socially deprived groups, those who are more likely to lack internet access include those who are retired, living in rural areas, unemployed, or less likely to live in a household with children (Helsper 2008).

A number of factors improve the meaning of information, such as clarifying information. Presenting risks in terms of frequency (eg, '1 in 1,000 people have x') rather than probability (eg, 'there's a 0.1 per cent chance of getting x') makes information more memorable and leads to different risk perceptions (Hibbard and Peters 2003). The way statistics are framed is also important. A gain frame (survival rate) is more effective with choices around prevention. A loss frame (mortality) is more effective to convey the effect of dying than a gain frame (survival rate) (Peters *et al* 2009). When indicators are partitioned in groups (Fox *et al* 2005), people pay more attention to the differences between partitions. For example, when information about takeaway meals is grouped by content (vegetables, pasta, meat), it makes consumers pay more attention to the nutritional content than when the same meals are presented grouped by cuisine (eg, Indian, Italian, Chinese). Other studies have shown that use of consistent labelling can increase comprehension and promote better decision-making. For example, providing non-numeric labels such as 'fair' or 'excellent' can translate the meaning of numerical information for those who are less numerate. Pairing numerical performance information with non-numerical labels was found to increase use of information by less numerate people in experiments on US health plan choices (Peters *et al* 2009).

In addition to strategies which seek to improve the use of comparative information by decreasing the quantity of information or changing the visual display, there are other ways to enhance patients' use of information on health websites. This can involve addressing the level of information offered, lowering the cognitive effort required, and utilising decision tools. Suggestions to enhance consumer use of information include:

■ increasing computer-aided decision tools

■ having an information intermediary

- helping people have a better idea of what actual experience of choice might be, as many have no experience of making these choices

- having narratives about someone's experience of choosing (however, as previously mentioned, this can bias decision-making)

- framing (for example, the effect of dying seems greater when framed as mortality rather than survival rate, and a gain frame is more effective with choices around prevention) (Hibbard and Peters 2003)

- putting the most important information first in any list (Lohse 1997).

The London Choice Project (LCP) used two strategies to improve decisions: patient care advisers and decision aids. Patient care advisers reduced the variations in uptake of choice according to social class, education level and ethnic group (Ellins and McIver 2009).

Using nudges to help people make the best choice

Nudges (Thaler and Sunstein 2008) have become a popular topic in current policy debates to improve decision-making. They have been shown to have a particularly strong influence on decisions in experimental research. Behavioural research has repeatedly established that preferences are constructed by the task (*see*, for example, Lichtenstein and Slovic 2006). Choice processes are informed by, and adapt to, the format of the information seen (Bettman *et al* 1993). This propensity to construct and adapt decision-making to the presentation given creates opportunities for 'nudging' participants towards better choices, by implementing specific nudges.

A detailed description of a number of nudges can be found in Thaler and Sunstein (2008). The box opposite summarises a range of diverse tools that have been suggested and, in principle, could be used to improve decision-making, with examples of how they are used in other areas of public policy.

There is a great deal of discussion about the ethics, feasibility and effectiveness of using nudges in policy-making, but we are not aware of any studies that have contrasted the effectiveness of these different nudges on patient decision-making. In particular, most of the current discussion seems to revolve around financial nudges (incentives) or using defaults and structuring of complex information. There appears to be less interest in the areas of mapping, feedback and error. This research gave us the opportunity to contrast the effectiveness of different nudges and to test how susceptible different demographic groups are to different nudges, including those that appear to have been neglected so far. In the context of hospital choice, we were particularly interested in the use of nudges to help patients avoid hospitals with poor clinical quality and safety records, and to choose the highest quality provider. However, policy-makers may have other objectives – to nudge patients to make decisions in line with their preferences or to minimise anxiety in making a choice.

Definitions and examples of NUDGES

Financial i**N**centives: financial losses or gains which seek to influence decisions

- Entering people into a lottery if they sign up to a weight loss programme

- Destiny Health Plan: Vitality Bucks to make healthy choices (gym, blood pressure check)

- Dollar a Day: teenage girls with a baby receive a dollar for each day they are not pregnant

Understanding mappings: tools which help people to understand and define preferences better

- Decision support tools for prostate cancer patients to weigh up the risks and benefits of treatment options, based on relative value of urinary and sexual function

- RECAP (Record, Evaluate, Compare Alternative Prices) for credit cards, mortgages, insurance policies.

Default: preselect the desirable option and require people to opt out

- Pension plans set the default to be in the plan, requires people to actively opt out

- organ donation in some countries

- automatic renewal for magazine subscription

Giving feedback: offer feedback on the choice, selection or decision made and offer the opportunity to correct

- Digital cameras allow the user to see the picture after every shot

- ceiling paint that goes pink when wet but turns white when dry

- red light when filter for air conditioners needs to be changed

Expecting error: expect people to make mistakes

- Requirement for sellers of insurance products to allow a cooling off period in which people can change their mind

- Warning systems in cars (e.g. if you don't buckle your seat belt)

- Birth control pills that are taken every day without break

- 'Look right!' signs in London

- Gambling self-ban

Structuring complex choices: present information so it is easier to evaluate

- Key features of financial services products must be set out in a standardised format to make products easier to compare

- Colour samples in a paint store ordered by similarity.

Adapted from Thaler and Sunstein (2008, p 100).

Conclusion

Hospital choices are, per se, 'unfriendly' (Schlesinger 2010), as there is a lot of information on a number of important criteria, most of which are typically conflicting; for example, a patient may choose a hospital that is close to their home, but with low overall quality, or one that has higher than average risk, but where patients report that they are treated with more respect. In unfriendly environments like these, research has already found that people want to overcome their natural tendency to resort to simple heuristics that avoid trade-offs (Payne *et al* 1993; Gigerenzer *et al* 1999; Shanteau and Thomas 2000) and shift to more compensatory, information-intensive strategies that try to face trade-offs head on (Bettman *et al* 1993; Carrigan *et al* 2004). The result of this shift is effortful and emotionally difficult choices (see, for example, Luce *et al* 1999). Further, not everyone is able to make this shift, especially older adults (Mata *et al* 2007), who tend to stick to simpler strategies. Last but not least, health care choices are more emotional than regular consumer choices, and research has established that when choices are emotional, even people with high cognitive ability resort to simpler, less accurate heuristics (Luce *et al* 1999). This leads us to expect that only a small proportion of patients will be able to make effective decisions and cope in this unfriendly choice environment.

Despite some interest in applying the lessons from decision research to the design of comparative hospital reports, there have been few attempts to fundamentally change how they are presented. If the use of publicly reported quality information is to increase in the future, then the information presented on websites such as NHS Choices needs to improve to be both relevant and accessible.

2 Methods

Study design

We conducted a small-scale, qualitative study, comprising seven focus groups held across London. The results from the focus groups informed the design of a large-scale, two-phase experimental study that tested online hospital scorecards. The methods used and the findings from the focus groups are reported elsewhere in detail (*see* Fasolo *et al* 2010). Here, we present the findings of the online experiment, and draw on the focus group findings as appropriate.

The key aim of the online experiment was to compare how different scorecard designs affected the choices people made, identifying which designs worked best, and for whom. We sought to understand whether participants felt they had made an informed choice, whether they felt they had enough time to make decisions, and whether they were confident about and satisfied with their choices. We also measured the extent to which the effect of nudges lasts over time by testing at two different points in time. We were interested in whether the nudges helped to create consumers who are better informed, rather than just influencing an immediate choice.

Scorecard information and design

The focus groups helped to inform the design of the two-phase online experiment. Compared to NHS Choices, our online scorecard had fewer indicators, with clearer and more consistent definitions, simpler and more consistent formatting of data, and a balance in the number of indicators in each of the three domains of quality (clinical effectiveness, safety and patient experience) as well as location.

Selecting indicators

Currently, NHS Choices includes 36 indicators (*see* box on page 18–19) organised under 15 drop-down headings as part of the 'compare hospitals' pages. On other parts of the site, it publishes data on more than 500 indicators on topics like cancer and long-term conditions.

Indicators currently available on NHS Choices for hip replacement

Number of operations performed

- Number of hip replacement operations performed last year

Waiting times

- Percentage of patients seen within 19 weeks from GP referral to treatment

- Average time from GP referral to treatment

Average time spent in hospital

- Average time spent in hospital (days)

Patient outcomes for this operation at these hospitals

- Rate of unplanned readmissions (as expected, higher or lower than expected)

- Survival rates (within the expected range)

Levels of infection at these hospitals

- Wound infection rate (similar to the national average)

- Number of weeks MRSA-free

- Number of cases of MRSA within the past 12 weeks

Overall hospital performance

- Overall quality score (weak, good, excellent, with coloured spots)

- Hospital Standardised Mortality Rate (HSMR) (that is, the ratio of actual to expected deaths) (lower than, similar to or higher than the national average)

What patients say about their overall care at these hospitals

- Patient survey score for overall care for inpatients (out of 10)

- Number of people who rated the hospital on NHS Choices and would recommend the hospital to a friend

What visitors to NHS Choices say about these hospitals

- How clean were the treatment areas of the hospital? (clean, very clean) (power bar up to 5 squares)

- Patients said doctors and nurses worked well together... (some of the time, most of the time) (power bar up to 5 squares)

- Patients were treated with dignity and respect... (power bar up to 5 squares)

- Patients felt they were involved in their care... (power bar up to 5 squares)

Indicators currently available on NHS Choices for hip replacement *continued*

Results from the national survey of inpatients

- Patient survey score for dignity and respect (out of 10)

- Patient survey score for involvement in decisions (out of 10)

- Patient survey score for cleanliness of wards (out of 10)

- Patient survey score for availability of same-sex accommodation (out of 10)

Results from NHS staff survey

- Staff survey score for staff feeling satisfied with the quality of work and patient care they are able to deliver (out of 10)

- Staff survey score for staff receiving job-relevant training, learning or development in the past 12 months (out of 10)

- Staff survey score for staff agreeing that, if a friend or relative needed treatment, they would be happy with the standard of care provided (out of 10)

- Staff survey score for staff that would recommend their trust as a place to work (out of 10)

Reporting of patient safety incidents

- Number of patient safety incidents reported to the National Reporting and Learning System (x out of the past six months)

Quality of food provided at these hospitals

- Overall quality of the food (acceptable, good, excellent)

Car parking

- Car parking (yes/no)

- Number of car parking spaces

- Disabled parking (yes/no)

- Number of disabled parking spaces

- Average hourly cost of parking (£)

Disabled access

- Disabled access (yes/no/data not available)

Induction loop, signing and translation services

- Induction loop (yes/data not available)

- Signing services (yes/data not available)

- Translation services (yes/data not available)

The scorecard in our online experiment included the indicators that focus group participants had identified as important. When asked spontaneously in the focus groups what factors are important, people mentioned a wide range of things, but the experience and expertise of particular doctors was cited most frequently. As this information is not yet routinely available, we did not include this in the experimental scorecard. However, when presented with a set of indicators about the quality of care, patients valued clinical aspects such as infection rates and survival rates. Nine indicators were organised into four domains:

- How good is the quality of care provided at the hospital? (subsequently referred to as quality of care)

- How safe is the care provided at the hospital? (safety)

- What do patients think about the hospital? (patient experience)

- Location.

These headings were used to organise the information. The wording used in the focus groups for specific indicators was taken directly from the NHS Choices website; this was simplified for inclusion in the scorecards used later on in the online experiment (*see* box below).

List of indicators used in the online experiment

How good is the quality of care provided at the hospital?

- What is the Hospital Standardised Mortality Rate (HSMR) at this hospital?

- How many people treated in this hospital reported an improvement in their health?

- What is the risk of having to return to hospital urgently within one month of a planned operation?

How safe is the care provided at the hospital?

- How many MRSA blood infections for elective patients are there in this hospital?

- How many people develop a wound infection after surgery at this hospital?

What do patients think about the hospital?

- How did inpatients score the level of dignity and respect shown to them?

- How did inpatients score the cleanliness of the treatment areas in the hospital?

- How did inpatients score their involvement in decisions about their care?

Location

- How far is the hospital from me?

To make values easy to evaluate, they were presented in numeric, verbal and graphic format (*see* Figure 4 opposite). We used a standard set of evaluative labels for each indicator: 'poor', 'good' and 'excellent'. These were used instead of more statistical but less meaningful terms such as 'above average' or 'below average', 'as expected', and so on. We used one, two or three ticks corresponding to the evaluative categories ('poor' = 1 tick, 'good' = 2 ticks, 'excellent' = 3 ticks).

Scorecard design

The basic design of the scorecard used in the experimental phase was based on the latest NHS Choices design. It changed significantly between the time when the focus groups took place and the start of the online experiment. The programming and design was done by Gravitas Research. The current NHS Choices version allows people to see the search results displayed as a simple list, a detailed comparison, or a map. We used the detailed comparison version, which displays five hospitals in columns across the screen (*see* Figure 4 below).

On the NHS Choices website, a photo of the hospital appears at the top of the column, with the address and the distance in miles. In our study, hospitals were not given names but were designated by letters (eg, Hospital S, Hospital W), with a photo of a generic hospital building (all modern).

On NHS Choices, the distance selected can be 5, 10, 20, 50 or 100 miles, or any distance. The default is 50 miles. In our study, we presented five hospitals in total, with a range of

Figure 4 Screenshot of the scorecard used in the online experiment
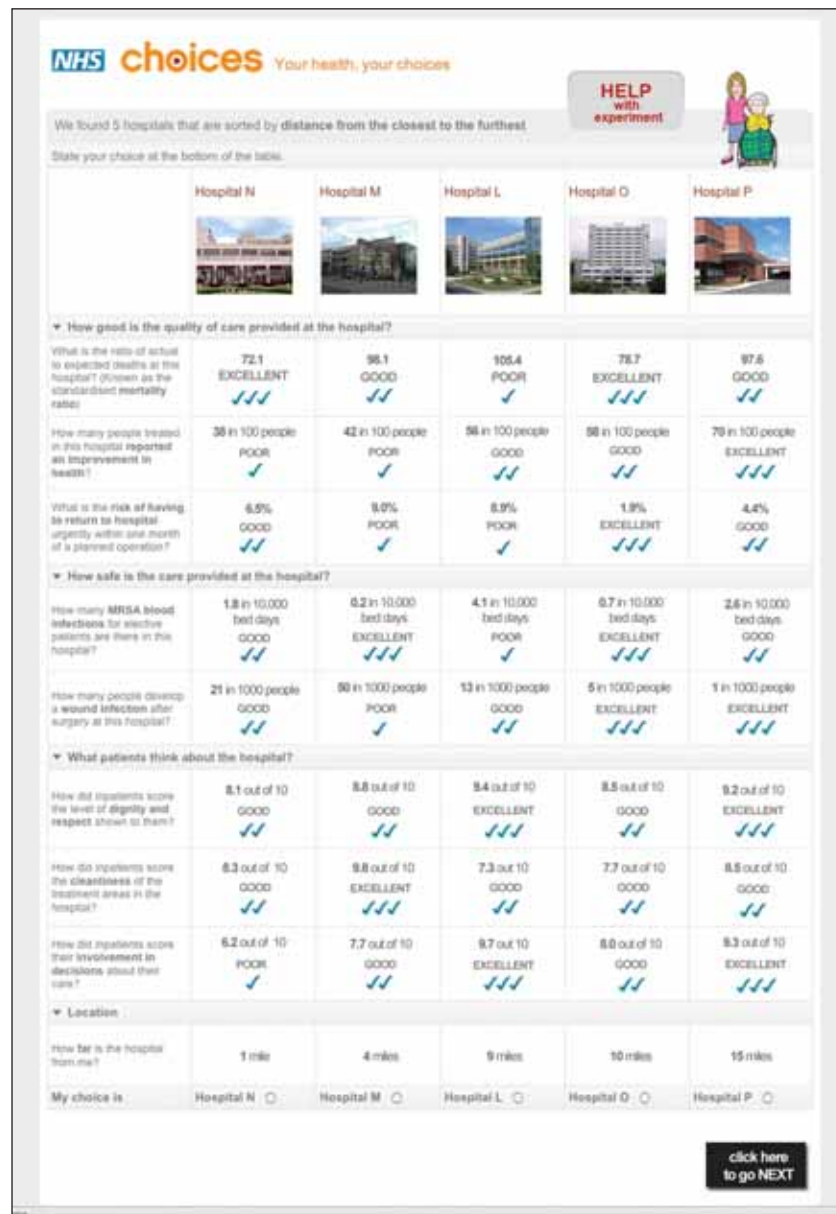
up to 18 miles maximum. The distance to the hospital was displayed at the bottom of the scorecard (as on NHS Choices).

On NHS Choices, the hospitals are ordered by distance, with the closest shown first (on the left of the screen). Alternatively, they can be presented alphabetically or, in a more recent feature, ordered by quality (lower than average wound infections, excellent overall quality, lower than average mortality and lower than average readmission rates shown first). We had two control scorecards: one in which the hospitals were sorted by distance (distance sort), from left to right, and the other in which the hospitals were sorted by quality (quality sort), from highest (left) to lowest (right).

To make the scorecards as realistic as possible, they were populated in such a way that no hospital was 'best' on all nine indicators. That is, if a hospital was best on one indicator (eg, had the lowest mortality ratio), it performed worst on other indicators (eg, was the furthest in terms of the 'location' indicator).

However, the values were chosen such that there was always a hospital that was best for clinical effectiveness indicators of quality of care (eg, had two 'excellent' clinical quality indicators and one 'good' one). This allowed us to capture the difficulty of trade-offs faced by patients in real life, while at the same time enabling us to identify a normative 'highest quality hospital' and therefore judge whether patients had actually chosen the highest quality hospital. Most previous research on choice has either been explicitly designed to examine trade-offs (for example, discrete choice experiments (Burge *et al* 2006)) or has used a design which has a clear 'best overall' hospital that patients are expected to agree on and choose (Hibbard *et al* 2007).

Using nudges to manipulate choice

Using Thaler and Sunstein's framework (*see* previous chapter), our research focuses on nudges that we judged to be applicable to the context of hospital choice: Defaults, Expect error, give Feedback and Understand mappings. Financial incentives did not seem applicable in the NHS context because it is difficult to imagine a situation where patients would be paid for choosing a high-quality hospital – although reduced co-payments are used in some US insurance plans if patients choose hospitals within a pre-selected network. We sought to integrate as many of the strategies to structure complex choices as possible in our scorecard (described above), but did not seek to assess their relative impact on choice.

We designed and employed three types of nudges in the online experiment, as follows.

■ Default (pre-select nudge): pre-selected the 'highest quality' hospital in terms of clinical effectiveness and quality of care for patients.

■ Understand mappings (preference nudge): provide tools which help people to better understand and define their preferences.

■ Expect error and give feedback (feedback nudge): offer feedback to people who based their choices on non-clinical indicators, and offer them the opportunity to reconsider their choice.

We included the default nudge because it is the most frequently discussed nudge, and has been adopted in a number of policy areas (for example, in relation to pensions). However, defaults have been mainly researched with regard to their ability to help with decisions to either do or not do something (*see*, for example, Johnson and Goldstein 2003). But they

have never been tested in the context of a multi-option, multi-criteria hospital choice ('choose 1 out of N'). We also included preference mapping in our experiment because it is particularly suited to problems characterised by multiple options and multi-criteria (Chernev 2003; Carlson and Bond 2006). Finally, we included feedback, as we considered that by making people aware of their mistakes and highlighting the salience of the information about clinical quality, it would create patients who were more informed.

Each of the nudges we used in the experiment involves different cognitive processes. The default nudge is a nudge for the 'lower brain' (Halpern 2009, p 233) while the other two (preference and feedback) are nudges that appeal more to the 'higher brain', which deliberates. In this report, we refer to these respectively as a 'mindless' nudge (default, or pre-select nudge) and 'mindful' nudges (preference and feedback).

The following seven scorecard designs were used in phase 1 of the online experiment. The first two are different versions of the current NHS Choices website.

Control scorecards

Distance sort This scorecard presented hospitals using the NHS Choices design and sorted them according to distance (listing the closest hospital first and displaying the text: 'We found five hospitals that are sorted by distance, from the closest to the furthest'). The highest quality hospital was in the fourth column.

Quality sort This scorecard presented hospitals using the NHS Choices design and sorted them by quality (listing the highest quality hospital first and displaying the text). 'We found five hospitals that are sorted by quality, from the best to the worst' (the highest quality hospital was in the first/left-hand column).

Nudge scorecards

(Mindless) pre-select nudge This scorecard presented information on the five hospitals, but one hospital was highlighted and pre-selected, so that this hospital was chosen unless the participant opted out (*see* Figure 5 overleaf). The default hospital was the 'highest quality' hospital in terms of clinical quality of care, and was presented first (in the left-hand column).

(Mindful) preference nudge This nudge is based on decision theory and, in particular, the notion of articulation of preference and importance of choice criteria. In this nudge, participants saw three additional pages before seeing the scorecard: the first page made participants aware that hospitals typically vary greatly along nine hospital indicators (ie, 'no hospital is excellent across all indicators') and that they would need to make trade-offs (*see* Appendix 1, online, available at www.kingsfund.org.uk/hospitalchoice). The second page explained the nine indicators, and the third page asked participants to rank them according to how important and relevant they felt each indicator was, 1 being 'most important' and 9 'least important'. This scorecard design gave participants the opportunity to consider all the indicators in the online experiment before choosing a hospital.

(Mindful) feedback nudge This scorecard included a prompt to reconsider choices when participants had not chosen the highest quality hospital. The prompt pointed to the 'sub-optimal' quality of the chosen hospital and gave participants the opportunity to continue or to 'reconsider their choice' – that is, make the choice again (*see* box on page 24).

Figure 5   Screenshot showing the pre-select nudge with the highest quality hospital highlighted in yellow
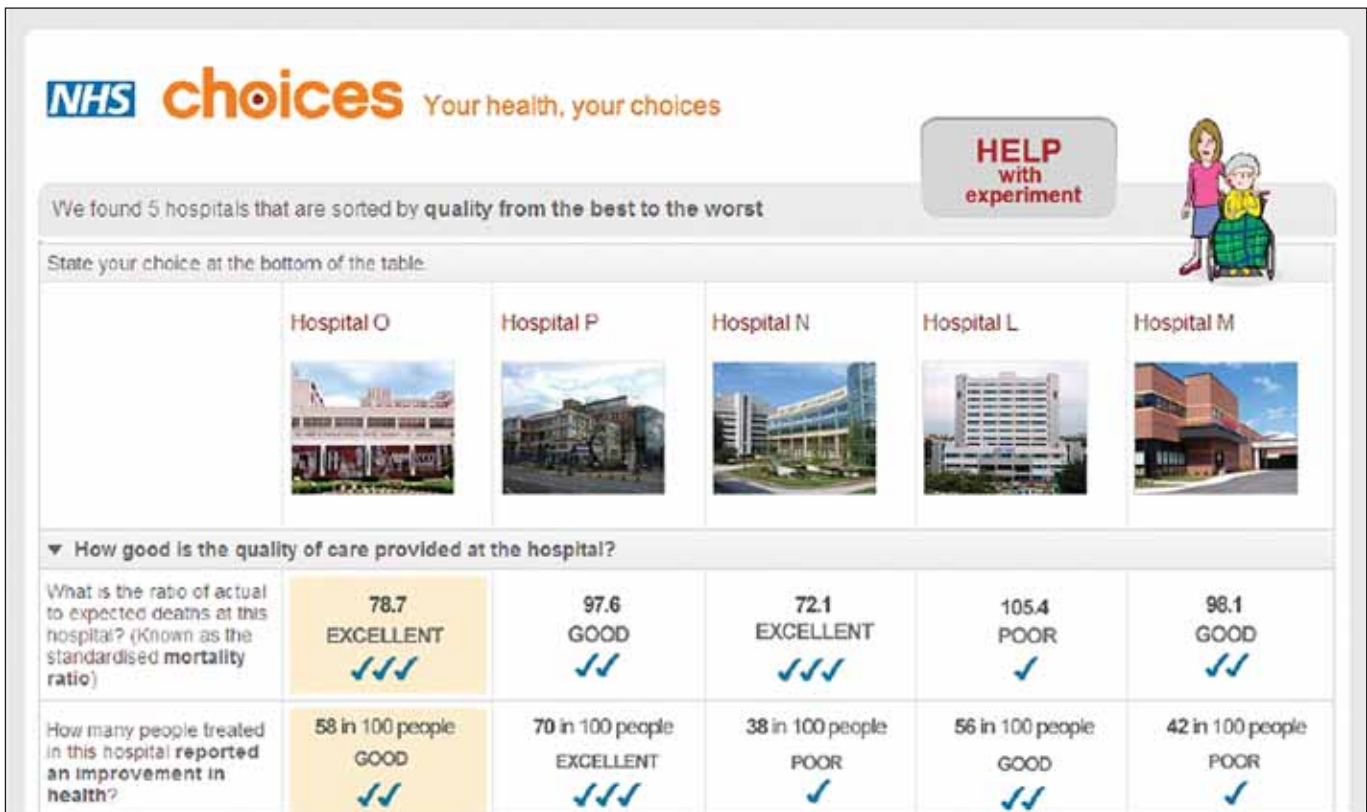
> **Details of the feedback nudge**
>
> If participants did not select the highest quality hospital, the following message popped up when they made their selection:
>
> 'Please confirm your choice. You have currently selected a hospital that, on the scale from 'poor' to 'excellent', has a 'poor' [or 'good'] quality rating. Are you sure you want to proceed and select this hospital, or would you like to reconsider your decision?'
>
> Participants then had to select either: 'I want to reconsider my selection' or 'I want to keep my selection'. If they chose to reconsider, they were presented with the same message until they had either chosen the highest quality hospital or selected to keep their second-rate selection.

Information controls

To ensure that any observed differences between the preference nudge scorecard and the control scorecards was not due to participants being given information explaining that hospitals varied greatly, two more scorecards were created:

Distance sort with information   This scorecard was identical to distance sort, but also included an extra page to inform participants that hospitals vary greatly on several indicators.

Quality sort with information   This scorecard was identical to quality sort, but also included an extra page to inform participants that hospitals vary greatly on several indicators.

Online experiment

There were two phases to the online experiment.

In phase 1, participants were asked to make two hospital choices and one hotel choice (not reported here).[1] For each choice, participants were presented with a slightly different scenario – they had to imagine they needed treatment for a non-urgent knee problem (hospital choice 1, or C1) and a non-urgent cataract (hospital choice 2, or C2). Participants were randomly assigned to one of the seven scorecard designs detailed above.

In phase 2 (several weeks later), participants were asked to make one hospital choice (C3) and one hotel choice (not reported here). Only two scorecards were used – distance sort or quality sort. None of the nudge scorecards were used in phase 2. Participants had to imagine having a serious non-urgent gallbladder problem for which they needed surgery. The indicators used were the same for each choice, but there were different values used in the scorecard.

The experiments included a series of questions after each choice regarding the participants' satisfaction with the choice made, confidence in their decision, and the level of difficulty in making a choice. Participants were also asked factual questions to test their comprehension and awareness of the information presented, and were asked about the extent to which they felt informed. Each respondent was asked to give basic demographic information, as well as information about how they use the internet and health services. We also asked them to respond to a short-form numeracy test consisting of three questions (adapted from Lipkus *et al* 2001). See Appendix 2 (online, available at www.kingsfund.org.uk/hospitalchoice) for full details of the questions asked.

In order to find out how participants used the information on the scorecard, and how long they spent looking at the different indicators, we used a technique called process-tracing. This is well known in consumer and marketing academic research, which 'traces' the decision process that consumers follow when deciding which product to buy (eg, Mouselab's research by Payne *et al* 1993; web-based Mouselab in Fasolo *et al* 2005). To allow process-tracing, the different options available are usually displayed in a scorecard (with products along columns, and features along rows – or vice versa), but the values in each cell of the scorecard are 'hidden' or blurred. To make the values of any feature or product visible, consumers have to point the mouse to the relevant cell. Then, its value is revealed and the values contained in any other part of the matrix are hidden. This emphasis on 'unobtrusive' tracing of how information is used during the actual process of choosing a hospital is novel and different from other research, which is usually limited to pre- and post-choice attitudes and responses, without any actual record of how people use the information provided when choosing a hospital.

The values of different indicators were blurred in the first and third hospital choice (*see* Figure 6 overleaf). Indicator labels and hospital names were always visible to participants. They were 'revealed' in the second hospital choice.

---

[1] We included a hotel choice as something people are more familiar with, and where the information is more easily understood. We were interested in comparing how people choose and the performance of the nudges for hotel choices with hospital choices. This analysis will be published elsewhere.

Figure 6 Scorecard with indicator information blurred
(hospital choices C1 and C3)
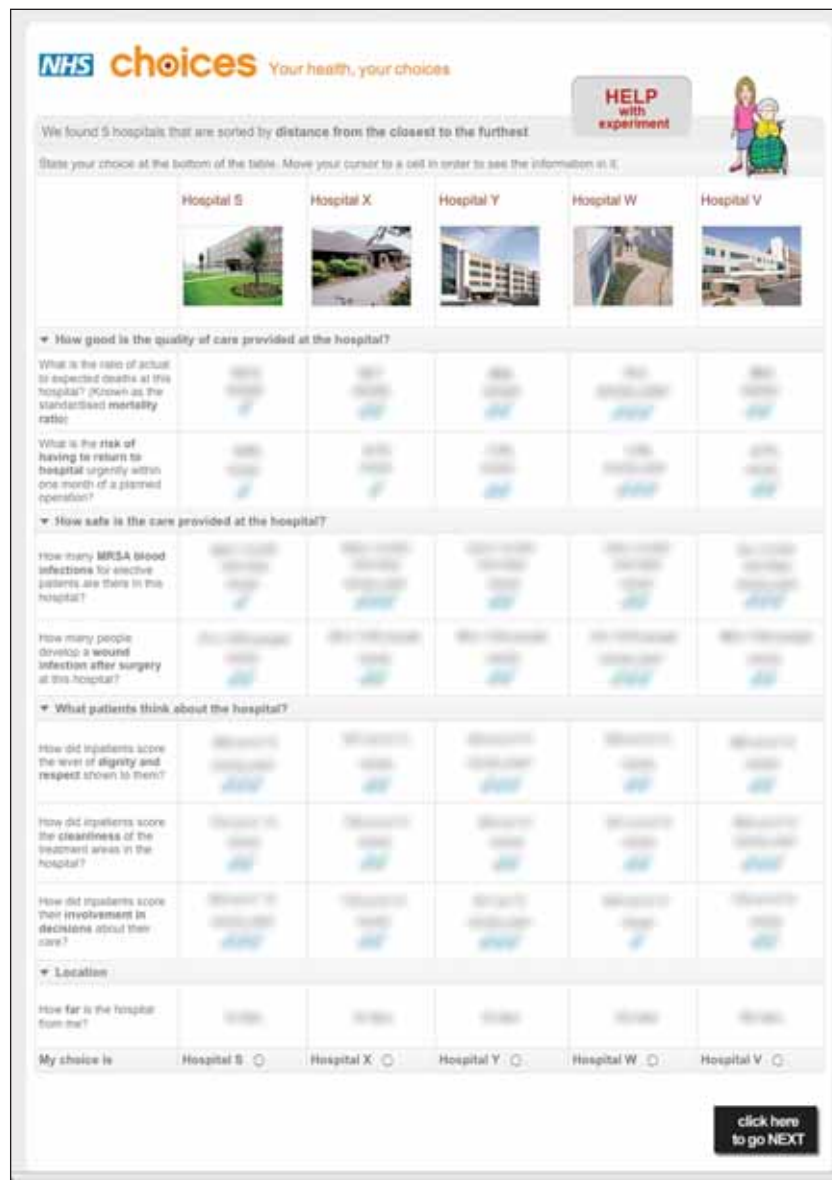
Table 1 Summary of hospital choices presented to participants

	Hospital choice 1 (C1)	Hospital choice 2 (C2)	Hospital choice 3 (C3)
Timing	First choice offered	Straight after C1	Two weeks after C1 and C2
Study phase	Phase 1	Phase 1	Phase 2
Scenario	'Imagine you have a serious non-urgent knee problem. Your GP recommends that you undergo knee replacement surgery. Choose the hospital where you would go.'	'Imagine you have a serious non-urgent cataract problem. Your GP recommends that you undergo cataract surgery. Choose the hospital where you would go.'	'Imagine you have a serious non-urgent gallbladder problem. Your GP recommends that you undergo surgery. Choose the hospital where you would go.'
Information visibility	Information blurred – made visible on mouse click	All information visible	Information blurred – made visible on mouse click
Scorecard designs	Randomly assigned to one of seven designs	Same as in C1	Randomly assigned to either quality sort or distance sort

Sample

The online experiment was sent to a stratified sample of English residents recruited through an online research panel that is representative of the English population. The online survey was conducted by Maximiles. Participants received loyalty points in proportion to the time spent completing the experiment, which they could redeem when shopping online.

Panel members were invited to participate in the first phase of the online experiment in January 2010, and in phase 2 in February 2010. Participants were informed before taking part in the study that the purpose of the experiment was to understand how patients search for information and make hospital choices. They were informed that they could participate if they were UK residents or had lived in the United Kingdom for more than five years and were familiar with the NHS.

For each scorecard, we aimed to have an approximately equal quota of young and old people, and people with high and low levels of education. In phase 1, the sample was randomly assigned to seven different scorecard designs (detailed on p 23). (We used a cut-off point for age and education, to give the dichotomous categories of young or old, and high or low level of education. 'Old' is defined as 51 or over, and 'young' is defined as 50 or under. 'High' education refers to those with at least some college education, whereas 'low' education refers to those with up to secondary education.)

An invitation to participate in phase 2 was sent to all those who completed phase 1 two to four weeks after doing so. In phase 2, participants were randomly assigned to either the distance sort or quality sort scorecards.

Data analysis

The responses of participants in both phases, across the different scorecard designs, were analysed and compared with regard to the following.

- The information people use when choosing a hospital.

- What people say influenced their decisions when choosing a hospital.

- How aware people were of differences in quality between hospitals, and trade-offs.

- What choices people made and who made the 'best' choices.

- How satisfied people felt when choosing a hospital.

- How scorecards can be improved to teach people to make better, more informed choices.

Process-tracing

Based on the data we obtained with mouse-tracing technology, we created 'heatmaps' of the different designs (*see* Figure 7 overleaf).

In eye-tracking research, heatmaps are a standard way of representing graphically which parts of a screen attract the most attention and which parts the least (*see*, for example, Reutskaja *et al* forthcoming).[2] The heatmap in Figure 7 is the graphical representation of the scorecard people saw, overlaid with mouse-tracing data and represented in colour. The results are separated according to the different scorecard designs, and concern the

---

[2] In previous research (Chandon *et al* 2007), eye fixations on certain locations were shown to be closely related to the attention people pay to a particular location (but not necessarily correlated with choice). We will build on this assumption when interpreting our results. Eye fixation and 'clicks' on the cells are not the same. However, we consider that 'clicks' are good approximations of eye fixations.

average time spent per cell and the total number of clicks per cell. On the heatmaps, cells where people spend the most time/click the most are 'hot' (red), while those where people spend the least time/click the least are 'cold' (green). The 'maximum' and 'minimum' time/number of clicks on the scale (which determines the colouring of cells) is determined for each map. Heatmaps help to understand how the data presented to participants in the form of the scorecards were acquired, evaluated and used. The number in each cell for 'Average time spent by participants per cell' represents the average amount of time spent per cell by people who clicked on this particular cell at least once.[3] The number in each cell for 'Total number of clicks per cell' represents the total number of clicks made by participants on this particular cell independent of the click duration (as long as it was longer than 250 milliseconds (ms)).

Figure 7 Example of heatmap

[3] Average time spent per cell is not affected by people who never clicked on this cell (those with 0 clicks), as in the computation of the average, we include only times greater than 0 – that is, people who clicked on the cell at least once.

Before creating the heatmaps, the process-tracing data-file was 'cleaned'. That is, we removed all the clicks that lasted less than 250ms (a quarter of a second) from the analysis. Eye-tracking literature suggests that 'eye fixation' usually lasts 200–250ms (Salvucci and Goldberg 2000; Reutskaja *et al* forthcoming). Any durations shorter than 250ms might have been just mouse movements through the cells and do not necessarily reflect the data acquisition process, so we deleted these observations.

Questionnaire

Responses to the questions in the online experiment were analysed using SPSS. A combination of descriptive statistics, correlations, bivariate and multivariate regression models were used to explore the relationship between the scorecard design, demographic characteristics, and choice. The main dependent variables of interest were choice of the highest quality hospital and satisfaction with the choice made. In each experiment, there was a choice of five hospitals. Binary variables were constructed based on the hospital chosen (choice of first-ranked hospital vs choice of any other hospital).

Satisfaction was measured on a 10-point scale, with 1 representing 'not at all satisfied', and 10 representing 'extremely satisfied'. We created a categorical variable, with those who rated 8 or above deemed 'very satisfied' and those who rated 7 or below simply 'satisfied'. Use of the internet was recoded into those who reported using the internet rarely (1–3 times per month) or regularly (1–3 times per week) [= rare to regular] and those who said they were 'always connected' or 'used the internet at least once a day' [= always connected].

Bivariate analysis was used to examine the extent to which different scorecard designs or demographic characteristics related to choice of hospital or satisfaction with choice. The Pearson chi-squared statistic was used to judge the level of association between factors and choice or satisfaction outcome.

To find out which particular scorecard designs or demographic characteristics are most associated with good hospital choices and satisfaction with choice, we ran a series of binary logistic regression models. Standard tests were run to establish the extent of collinearity between variables in the model.

The variables included were:

- age (18–34, 35–50, 51–65, 65+)

- gender (male or female)

- level of education (none, primary, secondary/high school, college, graduate/ postgraduate)

- internet use (rare to regular vs always connected)

- GP visits in the past year (none, 1–3, 4–6, 7+)

- hospital referral for the individual or a family member in the past year (none, 1–3, 4–6, 7+)

- previously offered a choice of hospital (never, 1–3, 4–6, 7+)

- seek health information online (very unlikely, very likely)

- numeracy (number of questions out of three answered correctly).

Due to the strong correlation between education and numeracy, we re-ran the model with education only and numeracy only. We present the results for numeracy and note any differences in the text.

Study participants

A total of 1,187 panel members completed the first phase experiment in more than 10 minutes and less than 95 minutes, and were therefore considered 'valid'. Anyone spending more or less time than this was deemed to not have taken the task seriously (ie, 90 people were excluded on these grounds). Of these 1,187 individuals, 19 were further excluded from the analyses as they either did not complete the experiment or did the experiment twice on different days. This left a total of 1,168 valid responses in phase 1. Completion times ranged from 20 to 30 minutes on average. Table 2, below, shows the characteristics of the participants in phase 1.

A total of 744 participants completed phase 2 in more than 5 minutes and less than 90 minutes,[4] and therefore were considered 'valid'. Of these, four were excluded from the analyses because they did the experiment twice or did not make any choice. Completion times averaged around 17 minutes. This means we lost 428 participants between phase 1 and phase 2. Table 2 also shows the characteristics of the participants in phase 2.

Table 2 Participants' characteristics across phases 1 and 2

Variables	Categories	Phase 1		Phase 2	
		N	%	N	%
Education	None	106	9	65	9
	Primary school	125	11	77	10
	Secondary school/high school	293	25	170	23
	Some college education, but not graduate	193	17	121	16
	Graduate/postgraduate	451	39	308	42
Age	18–34	491	42	301	41
	35–50	122	10	68	9
	51–65	470	40	308	42
	65+	85	7	64	9
Gender	Female	639	55	320	43
	Male	529	45	421	57
Numeracy scores	0	181	15	103	14
	1	277	24	172	23
	2	289	25	182	25
	3	421	36	284	38

[4] We made thresholds lower (ie, 5 and 90 minutes instead of 10 and 95) because the task was shorter in phase 2 than in phase 1.

3 Focus group findings

The first stage of our research explored how people understand and use indicators of hospital performance. The key findings from the focus groups are presented here. For a more detailed analysis, see Fasolo *et al* 2010.

What information do people want when choosing a hospital?

In each of the seven focus groups, the factors that participants said were important changed during the course of the discussion. During each two-hour session, we asked participants to identify what they felt was important to know when choosing a hospital, at four different stages, with the following results.

Stage 1: During the open discussion, participants stated they would choose a hospital based on the quality of doctors, availability of specialists, and distance from the hospital.

Stage 2: When prompted by being given a set of 16 indicators on cards, participants prioritised waiting times, cleanliness, and whether patients were treated with respect and dignity.

Following this short discussion, participants were asked to reassess their rankings, and their preferences again changed, with waiting times, survival rates and the risk of MRSA infection ranked most important. This suggests that at least some participants had been influenced by the preferences of others.

Stage 3: After having been presented with a scorecard and asked to choose a hospital, waiting times and risk of MRSA infection remained highly ranked, but participants also selected 'overall quality of service' (a summative measure of the hospital's performance).

Figure 8, overleaf, shows how participants' preferences for information about the quality of hospitals changed during the different stages of the focus group discussions. When prompted with other types of information, participants neglected almost half of the indicators they had first considered important; this suggests that interest in more complex information on clinical quality can be stimulated by the mere inclusion of this information.

The type of information people want is currently not provided on websites like NHS Choices. Participants wanted specific data about the quality of a particular treatment, not only about the overall quality of a hospital. They wanted data on individual clinicians or specialties, not data that are aggregated for the entire hospital. Scorecard-specific quality measures of hospital care in the United States suggest that there is significant variation in quality within hospitals as well as between hospitals (Jha *et al* 2005), thus supporting people's preferences.

> *If I had to be treated for cancer, I would put these cards in a specific way, but if I had diabetes, I would put these cards in a different way.*
>
> (Focus group participant, north London community centre)

Figure 8 Popularity of indicators at different stages of the focus group discussion

* Indicators presented to participants in Stages 2 and 3
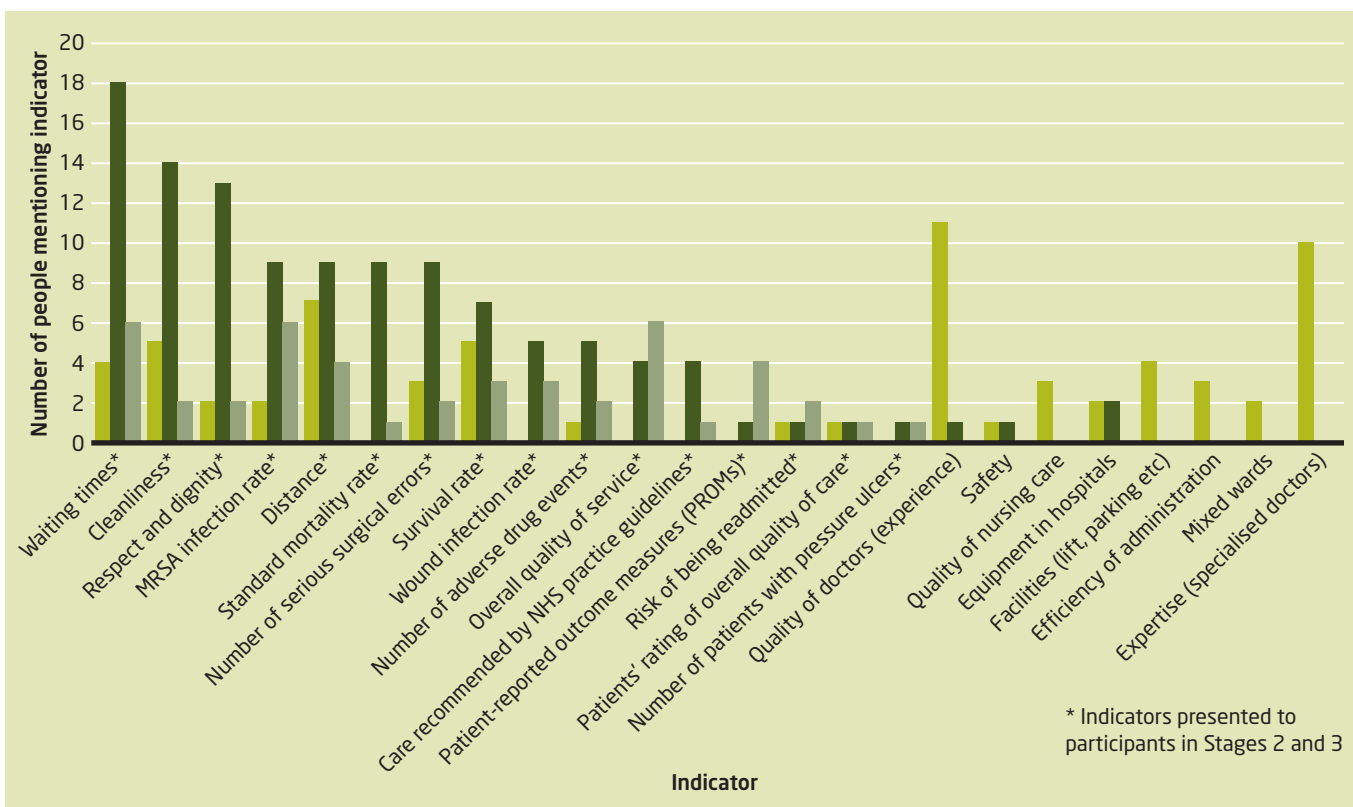
Stage 1
Stage 2
Stage 3

How do people use this information?

When presented with the 16 indicators described in Section 2, most participants thought they understood what the indicators meant. However, when asked to explain the indicators, their interpretation often differed from the definition and measures given on NHS Choices. Indicators were often misunderstood, and elderly and less educated patients had more doubts about their meaning, despite having more experience of hospitals. Even where patients do consult information, it is important to recognise that their choices may not be informed choices if the meaning of the information presented is poorly understood (*see* box on page 33).

The language used to describe each indicator influenced how well it was understood. Participants tended to rely on their initial understanding of each indicator when making decisions. More detailed explanations of indicators – for example, links to separate descriptions, or with notes alongside – were not given the same attention. It is vitally important to ensure that the initial, short indicator is easily understood.

When presented with the mock scorecard and asked to choose between hospitals, most participants tried to make sense of the whole scorecard rather than focusing on the indicators they had previously ranked as 'most important'. They tried to construct an overall assessment of the hospital, seeing it 'as a whole' rather than using their preferences to compare hospitals just on the indicator most important to them.

Participants' understanding of the Hospital Standardised Mortality Rate

The Hospital Standardised Mortality Rate (HSMR) is a complex measure. It measures the likelihood of a person admitted to hospital dying, compared to the general population in England. The number 100 represents that the area has an average mortality; a higher number means a hospital is performing poorly (higher than average mortality) and a low number means it is performing well (lower than average mortality). In the rest of the scorecard, high numbers signify that a hospital is doing well, so the HSMR is counter-intuitive to the rest of the figures. This is part of the reason why it is difficult to interpret.

I'm looking at what is the risk, and I'm not sure if a high score is good or bad here, for example... To me, it isn't intuitive. What's the risk? Is a high score a high risk or is it a low risk in fact?

(Participant, focus group 2, member of non-academic staff
at London School of Economics)

Most people understood the HSMR as a straightforward mortality rate:

What is the Hospital Standard[ised] Mortality Rate? People die, but how many do they save, and from the amount of people that go in, how many die? Maybe not ... their fault.

(Participant, focus group 3, north London community centre)

If you have a particular operation that you need done, if [it] has a good success rate... Unless it's a life or death question, you wouldn't look at mortality rates.

(Participant, focus group 4, member of non-academic staff
at the Institute of Education)

Instead of the risk level for the entire hospital, people wanted to compare and assess quality of care at specialty, team and procedure levels.

I'm looking (for) who'd operate on my heart. I want to know how many [they] did before, [how many] died in the process, how successful they are.

(Participant, focus group 2)

The way how I actually work through is, looking at each one of the criteria, on the balance of how important, and select what to exclude on the way.

(Participant, focus group 2)

I ticked which one I would choose for each particular question, and at the end, total up the amount of ticks, and number 3 has more ticks than the rest.

(Participant, focus group 1)

In addition to seeing data for the whole hospital, participants also tried to combine indicators themselves by detecting 'patterns' (eg, 'this hospital is mostly good' or 'mostly bad').

Participants found the inclusion of both aggregate and disaggregated information confusing. Summative indicators were popular with some, but others were unsure

what they were actually measuring. In the focus groups, the indicator that was most misunderstood was 'overall quality of service' – a measure based on the regulator's assessment of the hospital trust, and currently included on the NHS Choices website. Participants struggled to understand this category and interpreted it to be a general description of care in a hospital, covering nursing care, quality of the doctors, cleanliness, and control of infections. None of the participants understood the difference between 'overall quality of service' and 'how patients rate their overall care' – a different measure based on patient survey data.

> Participant 4 (P4): *My least important was 'what was the overall quality of service?'*
> Facilitator: *Why is that?*
> P4: *I ran out of numbers, basically.*
> Facilitator: *Does that mean that overall quality of service is not important?*
> P4: *Not really.*
> Facilitator: *What do you think this is?*
> P4: *The general way the hospital was run, its efficiency.*
> P1: *Whether they keep your records, your tests.*
> P4: *If they appear to be doing everything right.*
>
> (Discussion between participant and facilitator, focus group 7)

The appeal of aggregate measures differed according to demographic characteristics. Aggregate measures were more frequently used as a shortcut by older/less numerate participants, who found it harder to process the disaggregated information. Summative measures reduced the cognitive processing required by some individuals. In contrast, summative measures were considered too vague by younger/more numerate/higher educated participants. It is unlikely that only one format can meet the needs of a range of people.

Participants were confused where there were apparent inconsistencies between the summative score and individual items (eg, a hospital with a generally weak pattern of indicators but an 'excellent' score for 'overall quality of service'), which prompted some negative reactions. Different participants resolved these inconsistencies in different ways. Less educated participants (including those who were less numerate and older) tended to discount the inconsistent indicators, or worse, misinterpreted what other indicators meant (eg, a high number of MRSA cases must not mean high risk, if it is a hospital rated as 'excellent' overall or with a green traffic light).

Presenting information for each hospital side by side on all indicators in one single table led participants to realise that no one hospital can have top values across all indicators. Presentation of conflicting indicators forced people to think about making trade-offs. Many of the participants made trade-offs by 'hospital' (eg, does a better result in hospital 1 on indicator X compensate for a worse result on indicator Y?).

> P2: *The 'excellent' had high rates, but in other things...*
> P2: *I would want the 'excellent' one but it's the furthest from me...*
> P4: *But high number of drug events...*
> P1: *There is always a compromise, isn't there?*
>
> (Discussion between participants, focus group 7)

Making people think about these trade-offs and how much weight they should give to different factors (eg, going to a higher quality provider further away, or a lower quality provider nearby) is difficult, and can lead to dissatisfaction with choices. Participants did not express regret; they said they found it difficult to make trade-offs, but they did like having comparative information (Beattie and Barlas 2000; Spranca 2000).

As this information is confusing, people rely on other factors to help them make decisions. For example, personal experience and memories are a significant influence on how patients choose a hospital. Many focus group participants discussed memories of previous hospital experiences, and objective information on quality frequently failed to over-ride the importance of these personal views and memories.

The importance of their GP's opinion was often mentioned as influencing the choices people would make. People wanted their GP's opinion, in addition and sometimes instead of objective information provided by websites such as NHS Choices.

> *Unless you had some experience with that hospital or that particular element... you wouldn't know who is a good throat specialist in Charing Cross. So you can ask the doctor who he thinks, which is a good chance they'll know what doctors are practising or surgeons who are practising in the hospitals. So you still end up asking the doctor which one to go to.*
>
> (Participant, focus group 3)

How scorecard design influences people's choices

As discussed in 'Previous research' on page 10, subtle variations in the design of the scorecard can substantially alter people's preferences. For example, the order of presentation altered the attention people paid to different indicators; in our focus groups, waiting times were considered more important when presented first in the table than when presented seventh.

The scorecard used in the focus groups was similar to that used on the NHS Choices website; different indicators had mixed labelling formats – for example, ratios, scales, percentages and symbols. This made comparisons more difficult for participants. Many participants expressed the desire to have all the indicators presented using the same scale, symbol or colour scheme.

> *Any one of these systems would be fine, but having lots of different ones is confusing... There's a lot of information here, people would rather look at something straight away: 'Oh, green means good, or red means bad'.*
>
> (Participant, focus group 4)

The way numbers and symbols were represented also influenced comprehension and, ultimately, choice. Measurements using averages were considered more meaningful when given alongside ranges of values (eg, minimum and maximum value for a particular indicator). Colour schemes such as traffic lights were considered helpful by most (some wanted all the information colour-coded, while others wanted only key indicators colour-coded so that they stood out). Stars appealed mainly to elderly participants.

The use of numbers in indicators appealed to more numerate participants, but they wanted relative numbers that could be meaningfully compared. Again, inconsistency in indicators confused participants. There was misunderstanding of some indicators where high numbers signified something as a problem (eg, risk of developing a wound infection) when other indicators used high numbers to represent 'good' performance.

In addition, missing data attracted attention and generated negative reactions. Participants were highly suspicious when indicators stated 'data not available'. Many participants ignored indicators with missing information and did not use them to compare hospitals; others used the information only for the hospitals where the data were available.

Discussion

When asked spontaneously what factors are important when choosing a hospital, participants mentioned a wide range of things, but the experience and expertise of particular doctors was cited most frequently. As this information is not yet routinely available, we did not include this in the experimental scorecard. However, when presented with a set of indicators about the quality of care, patients valued clinical aspects such as infection rates and survival rates highly. The scorecard in the online experiment, therefore, included the indicators that focus group participants had, with hindsight, identified as important.

The findings from our focus groups confirm the findings of earlier experimental research. We identified a number of key guidelines which we used to inform the design of the scorecard for the online experiment.

Scorecard information used in the online experiment

- Use clear labels for indicators so that they can be easily understood without having to click through to further information.

- Use symbols and labels consistently.

- Where possible, simplify labels and make them consistent.

- Order indicators so that the most important are listed first.

- Present data on hospitals, not organisations.

- Where possible, make the data relevant to the procedure the patient is having.

- Do not include summative measures.

- Group indicators into domains.

- Apply evaluative labels (eg, 'poor', 'good', 'excellent').

Furthermore, the finding that preferences are unstable and that people are highly susceptible to how information is presented suggested that there could be an opportunity to influence or 'nudge' people to consider information that they might not have spontaneously thought important. We also found clear differences in information preferences, and ability to interpret the information and make a decision by age and education. We therefore sought to evaluate the impact of different nudges on patients, and whether these impacts differed by age, education and other demographic characteristics. Having set out the findings of the focus groups and how these influenced the design of the online experiment, we now turn to the findings of the online experiment itself.

4 Findings of the online experiment

What information did people use when choosing a hospital?

Information people used

- People spent most time looking at information presented first – that is, in the left-hand column and at the 'top'. Placing indicators at the top of the scorecard 'nudges' people to pay more attention to these.

- While people were able to compare information in any way they wanted – hospital by hospital (ie, by columns) or indicator by indicator (ie, by rows) – they spent more time viewing the first column (indicators for the first hospital) rather than the first row (how different hospitals perform on that one indicator).

- People did not necessarily choose the hospital they spend most time looking at; however, the more times people looked at individual indicators the more likely they were to choose the highest quality hospital.

- Those who were asked to consider what was most important to them first, acquired and used information more systematically.

- Mindful nudges prompted people to pay more systematic attention to quality and safety information.

The focus groups revealed the sort of information people say they want when choosing a hospital. However, they also revealed that simply by presenting information, it is possible to influence what type of information people say they want or would use. This suggests that the indicators and information people say they want and think they will use are not always the same as what they will use in practice if presented with information.

In the online experiment, we were able to answer the question: 'What information do people actually use when choosing a hospital?' We used responses from participants in the online experiment to understand how people used the information presented on the scorecard, and how the design of the scorecard influenced the way people used information. We considered the time spent looking at different items of information, the pattern of how information was acquired, and how often participants looked at information.

Details of the scorecard designs are given on p 23.

Average time spent per cell

On average, across the scorecards, participants spent 2.7 minutes making hospital choice 1 (C1), 1.7 minutes making hospital choice 2 (C2) and 2.1 minutes making hospital choice 3 (C3) (*see* Table 1, p 26, for details of choices).

The average time spent by participants per cell in milliseconds (ms)[5] is not distributed evenly across the scorecard (*see* Appendix 3, online, available at www.kingsfund.org.uk/hospitalchoice). In almost all of the scorecard designs, participants spent most time looking at information that was presented first – that is, in the left-hand column. In addition, people spent most time looking at the 'top' part of the scorecard rather than the bottom part. This effect is more apparent by columns than by rows – in other words, people appeared to spend more time scanning down the screen in order to understand all the indicators for one hospital. This suggests a specific data-acquisition process. While people are able to compare information in any way they want – hospital by hospital (ie, by columns) or indicator by indicator (ie, by rows) – they spent more time viewing the first column (indicators for the first hospital) rather than the first row (how different hospitals perform on that one indicator). It would be interesting to compare this orientation with one where hospitals appear by row rather than column to see whether, as we might expect, this would lead people to compare hospitals across a single indicator rather than look at how good a hospital is overall.

There also seems to be an area of the screen (the lower middle section) that is viewed the least across all scorecard types. This appears to be a blindspot.

The time people spent on collecting and evaluating information about hospitals is not entirely correlated with choice of hospitals, but rather, reflects the attention people pay to different parts of the scorecard. For example, people spent more time looking at the 'closest hospital' in the distance sort scorecard, which is presented first, yet they more often chose the highest quality hospital, which was presented fourth. This result is consistent with previous research, and reflects the fact that attention is not necessarily correlated with choice (Chandon *et al* 2007). With most of the scorecard designs, people spent longer looking at the last hospital (last column) than the columns in the middle. This might be an indication of people trying to understand the organisation of the table (ie, confirm their intuition that if the first column is the 'best', the last should be the 'worst'). This attraction to the last column might also indicate the attractiveness of 'extreme' locations on the display, as reported in previous research (Valenzuela and Raghubir 2010).

This means that if the closest or highest quality hospital is presented first, people spent more time looking at this hospital than any other. Similarly, if quality and safety indicators are put at the top of the scorecard, as they were in our experiment, they are noticed more than indicators at the bottom of the page (which, in our experiment, were patient experience and location). As such, placing indicators at the top of the scorecard 'nudges' people to pay more attention to these.

We were also interested in understanding whether the design of the scorecard affected the average time spent on each cell. There is no prior research comparing process-tracing during choice with different formats, but based on the limited literature, we expected more time/clicks on quality/safety with mindful (preference and feedback) nudges than the mindless (pre-select) nudge, and more systematic search of information by participants who had mindful nudge scorecards. We expected those with mindless nudge scorecards to be less systematic, with more emphasis on location/patient experience. The indicators were displayed in the same manner across scorecard designs, so any differences

---

[5] This analysis was done for C1 only.

in the time spent looking at indicators should not be driven by whether they were easy to understand, but rather, by the salience/importance of different information as a result of the different designs.

This expectation was confirmed by the heatmaps: participants who had mindful nudge scorecards collected and evaluated information differently from those who had other scorecard designs. Those who had the preference nudge, where they were asked to consider what was most important to them first, acquired and used information more systematically. They distributed the time they spent on the scorecard more evenly among hospitals, and seemed to compare hospitals indicator by indicator (focusing their time on 'most important' indicators), suggesting that they made a more informed choice, in line with their preferences.

Moreover, time is not equally distributed among different categories of indicators: in the preference nudge scorecard, the quality and safety categories are viewed longer than the other two categories (patient experience and location) – that is, rows with quality and safety indicators are 'hot'. Similarly, in the feedback scorecard, time is more evenly spread among the three quality indicators than in other scorecard designs.

Different scorecards affected how long participants looked at parts of the scorecard. When there was no mindful nudge (that is, distance sort, quality sort and pre-select nudge), people spent most of their time 'looking' at certain *hospitals* along all the indicators rather than at certain *indicators* for different hospitals – that is, there was no discrimination regarding time spent across rows, as if people perceived each indicator as equally important. In contrast, mindful nudges showed people using their time across rows in a more 'informed' manner – that is, more time was spent looking at the quality indicators.[6]

Total number of clicks per cell

Findings for the number of clicks per cell were similar in some respects to the amount of time spent. For example, cells at the top of the scorecard were looked at more often than cells in the lower part (*see* Appendix 4, online, available at www.kingsfund.org.uk/hospitalchoice).

Heatmaps for the total number of clicks per cell also confirm the finding that scorecard design affects the way information is acquired, evaluated and used. In mindful nudge and distance sort scorecards, more cells are 'hot' in the quality and safety indicator categories than in pre-select nudge or quality sort scorecards. This suggests that mindful nudges are effective, and prompt people to pay more systematic attention to quality and safety information.

In contrast to the finding regarding average time spent per cell, the number of clicks per cell was associated with hospital choice. People who were given preference nudge, feedback nudge or distance sort scorecards also made better quality choices than those who had the pre-select nudge or quality sort scorecard (*see* below).

Also in contrast with the analysis of time spent on each cell, the total number of clicks does not appear to be influenced by the order in which hospitals are presented. Hospitals that are chosen most often are also looked at (or clicked on) more often than hospitals that are chosen less often. Therefore, putting the most important indicators at the top of the scorecard (which, in turn, leads people to spend more time looking at them and click on them more often) might not only nudge people to pay attention to these indicators, but also nudge them to use the indicators more often when making their choices.

---

[6] Time spent on the scorecard was measured from the first click until the choice was submitted. In the case of the feedback nudge, the time included time spent looking at indicators both before and after being given feedback, up until the final decision was submitted.

Finally, the maximum number of clicks in the preference nudge is smaller than that in any other scorecard. This could mean that participants using the preference nudge scorecard acquired information more selectively (and still made very good choices).[7] Therefore, a mindful scorecard design, such as the preference nudge, might help people to use their 'limited resources' (time and effort) in a more efficient manner.

What influenced people's decisions when choosing a hospital?

What influenced decisions

- Exposure to different scorecard designs impacted on what people said influenced their choice.

- Feedback and especially preference nudges prompt people to make a more 'mindful' search of information and they are more likely to say that they were influenced by clinical indicators.

- Scorecard design also made a significant difference to whether participants chose the highest quality hospital.

- Numeracy was the best predictor of choosing the highest quality hospital in phase 1.

We used the responses of participants who took part in both phases of the online experiment to understand what people say influenced their decisions *after* choosing a hospital, across all of the scorecard designs.

Other hypothetical studies such as discrete choice experiments assess how much weight people put on different factors rather than asking them explicitly what they thought was important. Surveys of patients who have recently made a choice of hospital suggest there has been a shift, with issues such as waiting times becoming less important and aspects of quality becoming more important, although this may be due to changes in the way the question was asked (Dixon *et al* 2010). However, very few of these patients have consulted performance information such as that available on NHS Choices, and therefore may not be fully aware of the different factors they could consider.

It is interesting to note that the factors participants considered to be important changed during the course of the focus group discussions; when presented with more information and asked to make a choice, participants ranked clinical indicators as important. In the online experiment, participants were asked to select three indicators (from a fixed list) which influenced their choice the most in both phase 1 and phase 2 (*see* Figure 9 opposite). Over the two phases, on average, 50 per cent of the participants reported that MRSA infections and the mortality ratio were important influences on their choice, 41 per cent selected cleanliness, and 42 per cent selected the risk of returning to hospital. In general, participants claimed to have been influenced by clinical indicators of quality and safety. Only 12 per cent said they were influenced by the location of the hospital. However, these responses varied by scorecard and demographic characteristics

---

[7] One might argue that the maximum number of clicks in the preference nudge scorecard is lower than the distance sort or quality sort scorecards due to a smaller number of participants in the former. Though there is definitely some truth in this (eg, there are also fewer clicks and fewer number of subjects in default and feedback scorecards compared with distance or quality sort scorecards), the lower number of subjects cannot explain the lower number of clicks by itself. We can discard this argument when comparing the preference nudge scorecard to default or feedback scorecards, where the number of participants is similar to those who had the preference nudge.

Figure 9 Percentage of participants who chose each factor as one of three factors which influenced their hospital choice (phase 1 and phase 2)
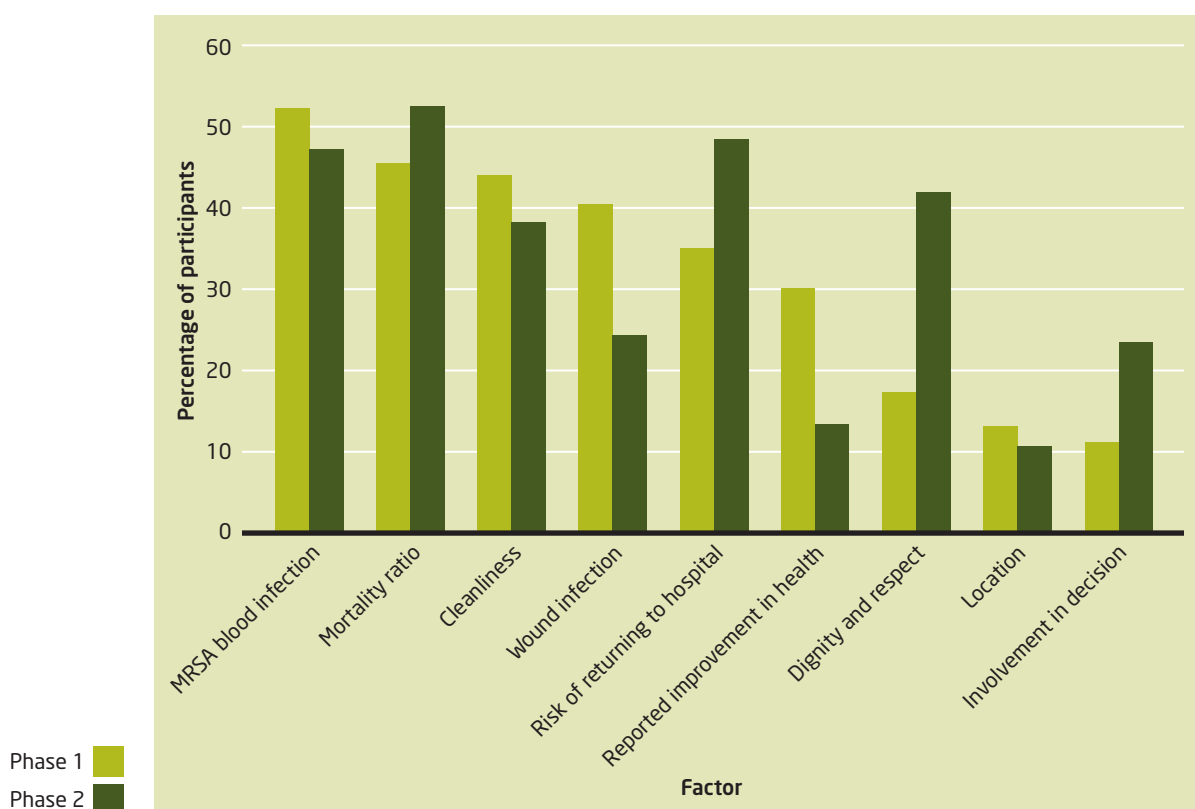
Phase 1
Phase 2

(*see* below). Between phase 1 and phase 2, there was a large increase in the percentage of participants choosing dignity and respect as a factor, and a large decrease in the percentage choosing health improvement as a factor.

Our findings are in line with recent surveys of real-life patients who have recently had to choose a hospital. Participants in the online experiment used values similar to those used in real life. Although the situation was hypothetical, it seems that having considered the information and been forced to make trade-offs, people value clinical factors. It would also seem that asking people what information they want without any context is not helpful in understanding what influences their choices.

We used the responses of phase 1 participants to understand the effect of different scorecard designs on what people said influenced their decisions. Exposure to different scorecard designs produced different responses. People who were exposed to any nudge – pre-select, preference or feedback – were significantly less influenced by cleanliness (41 per cent) compared with those who were not nudged (ie, those with distance and quality sort scorecards, 47 per cent).

Notably, participants who had the pre-select nudge scorecard reported being more influenced by non-clinical indicators than those with control scorecards (quality sort and distance sort). More specifically, participants using the pre-select nudge scorecard:

- were more likely to report having been influenced by whether patients were involved in decisions than those using the control scorecards (18 per cent vs 11 per cent)

- were less likely to report having been influenced by the mortality indicator than those using the control scorecards (38 per cent vs 46 per cent).

By using a pre-selected option, it appears that people are less influenced by clinical quality information than with the current NHS Choices design, in which hospitals are sorted by either distance or quality.

Of the two mindful nudges, the preference nudge appears to make the quality information more salient in people's minds than the feedback nudge.[8] Compared with participants who viewed the control scorecards (distance sort with information, and quality sort with information),[9] participants who had the preference nudge were also significantly less likely to cite cleanliness among the three most important indicators (20 per cent vs 50 per cent), and were significantly less likely to say they chose the hospital because it was closer (2 per cent vs 9 per cent).

Comparing the current controls with mindful nudges (preference and feedback) shows very similar effects. With mindful nudges, participants:

■ were less likely to cite cleanliness as one of the three most relevant indicators compared with those using the control scorecards (38 per cent vs 47 per cent)

■ were marginally less likely to cite involvement in the decision as one of the three most relevant indicators compared with those using the control scorecards (7 per cent vs 11 per cent).

This suggests that mindful nudges have the potential to influence the type of information people use to make decisions, and to give more weight to clinical information.

These findings reveal the factors that patients consciously say influenced their choice of hospital, whereas the process-tracing evidence tells us the indicators that 'unconsciously' influence choice. Combining both, we find that mindful nudges result in participants giving greater weight to clinical quality information. This evidence suggests that feedback and especially preference nudges prompt people to make a more 'mindful' search of information. These designs also make people more likely to say that they were influenced by clinical indicators (the information we sought to highlight in the experiment).

Awareness of differences in quality between hospitals

We used the responses of participants to understand the extent to which people are aware of differences in quality between hospitals, and of the need to make trade-offs between conflicting indicators such as high clinical quality and poor patient experience.

To find out whether participants were aware of the differences between hospitals, each was asked whether they agreed or disagreed with the following statements, using a 7-point scale where 1 = 'totally disagree' and 7 = 'totally agree'.

1. No hospital can be excellent on all indicators. A hospital that is poor on some indicators is bound to be good or excellent on others.

2. There is a big difference between a poor and a good hospital. I would never choose a hospital with a poor rating.

3. There is not much difference between an excellent and a good hospital. Both are good enough for me.

---

[8] There was no significant difference when directly comparing feedback and controls.

[9] When comparing with the control scorecards, we wanted to ensure that any difference would be due to the 'preference articulation and mapping of the indicators' and not to the fact that people were explicitly told that the hospitals varied on different indicators. For this reason, we compared the performance of the preference nudge against distance sort with information and quality sort with information (see 'Methods' section, pp 16–30).

Most participants (52 per cent) agreed that no hospital can be excellent across all dimensions of quality (*see* Figure 10 below). Seventy-seven per cent agreed that there is a big difference between a poor hospital and a good hospital, suggesting they would choose a good one over a poor one. Fifty-three per cent agreed that there is not much difference between an excellent hospital and a good hospital, suggesting they would not seek to maximise quality by seeking out an excellent provider. However, 26 per cent disagreed with this statement, suggesting they would choose an excellent provider over a good one. This suggests that people perceived a larger gap between the lower ranked and top ranked hospitals than between the highest quality hospital and the second ranked by quality, which would confirm the pattern of responses exhibited in Figure 10.

Figure 10 Levels of agreement with trade-off statements

Statement 1: 'No hospital can be excellent on all indicators. A hospital that is poor on some indicators is bound to be good or excellent on others.'

Statement 2: 'There is a big difference between a poor and a good hospital. I would never choose a hospital with a poor rating.'

Statement 3: 'There is not much difference between an excellent and a good hospital. Both are good enough for me.'
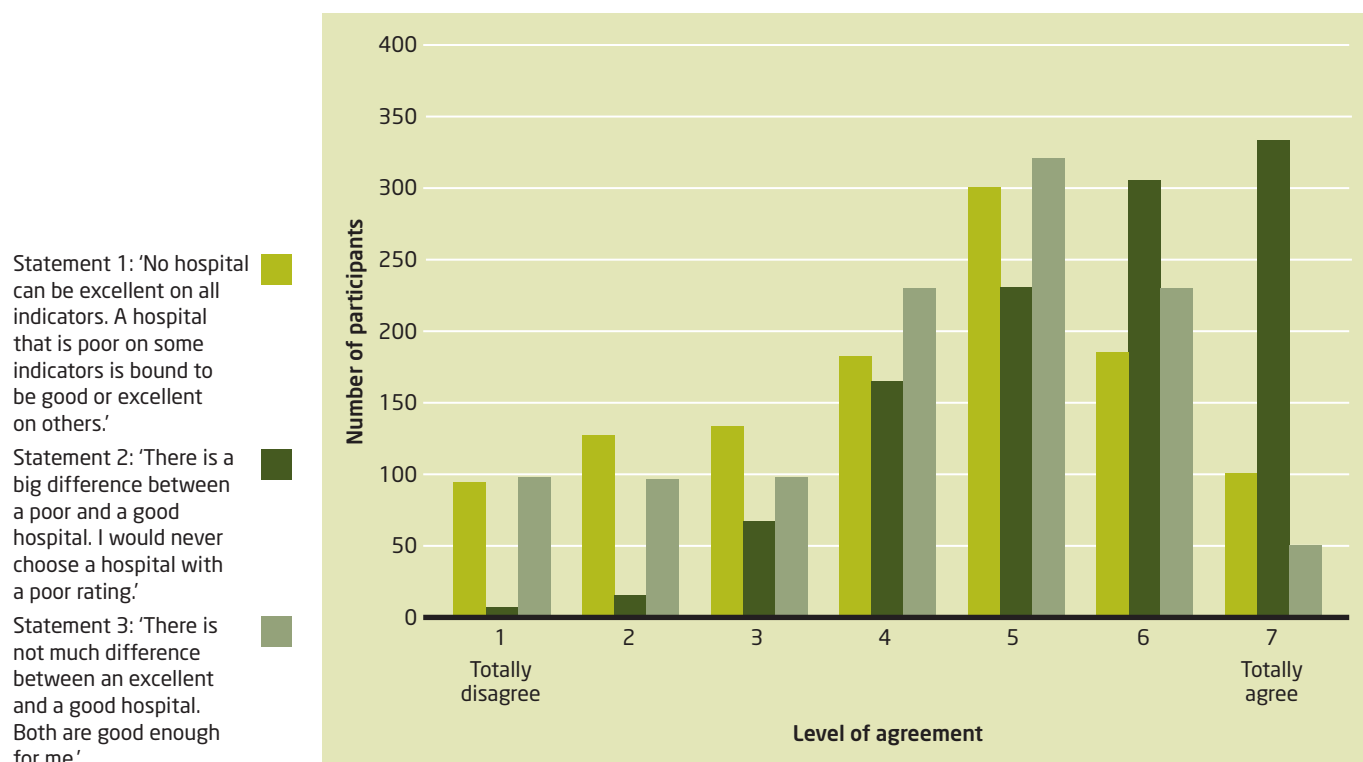
Participants were asked to rate the relative importance of different categories of indicators (quality of care, safety, patient experience, and location) via a series of pairwise comparisons on an 8-point scale. For instance, they were asked to state what (grouped as a pair) was more important to them (eg, quality of care or patient experience; safety or location). Answering these questions required participants to make a trade-off between the two categories (*see* Appendix 5, online, available at www.kingsfund.org.uk/hospitalchoice). However, participants were not consistent in how they responded to the series of comparisons, and these inconsistencies meant we could not construct a logical set of preferences for all respondents. Theoretically there were 24 possible logical hierarchies (of which all but one of these combinations was given by respondents). A third of respondents valued the categories in the following order: safety, quality, patient experience and location (*see* Figure 11 overleaf). Seventy per cent of respondents valued quality and safety above location and patient experience; 7 per cent of respondents' trade-offs resulted in illogical hierarchies (*see* Figure 12 overleaf).

This inconsistency of preferences is common, and decision support has developed in order to help people be more consistent. We did not use these responses to elicit rankings, but rather, to judge the relative weight given in the context of the pair.

Figure 11 Most common value hierarchies by percentage of respondents
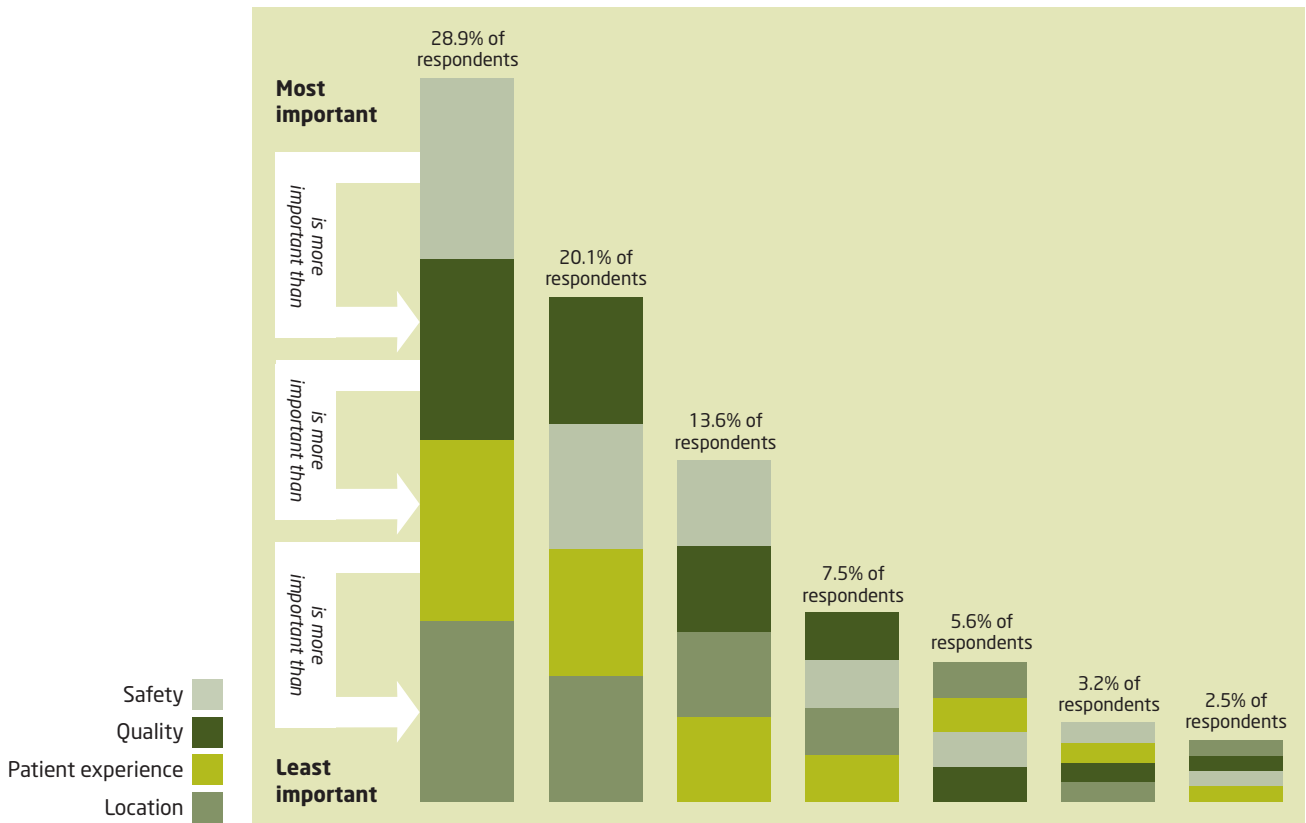
Figure 12 An example of an illogical hierarchy constructed from one response to the pairwise trade-offs
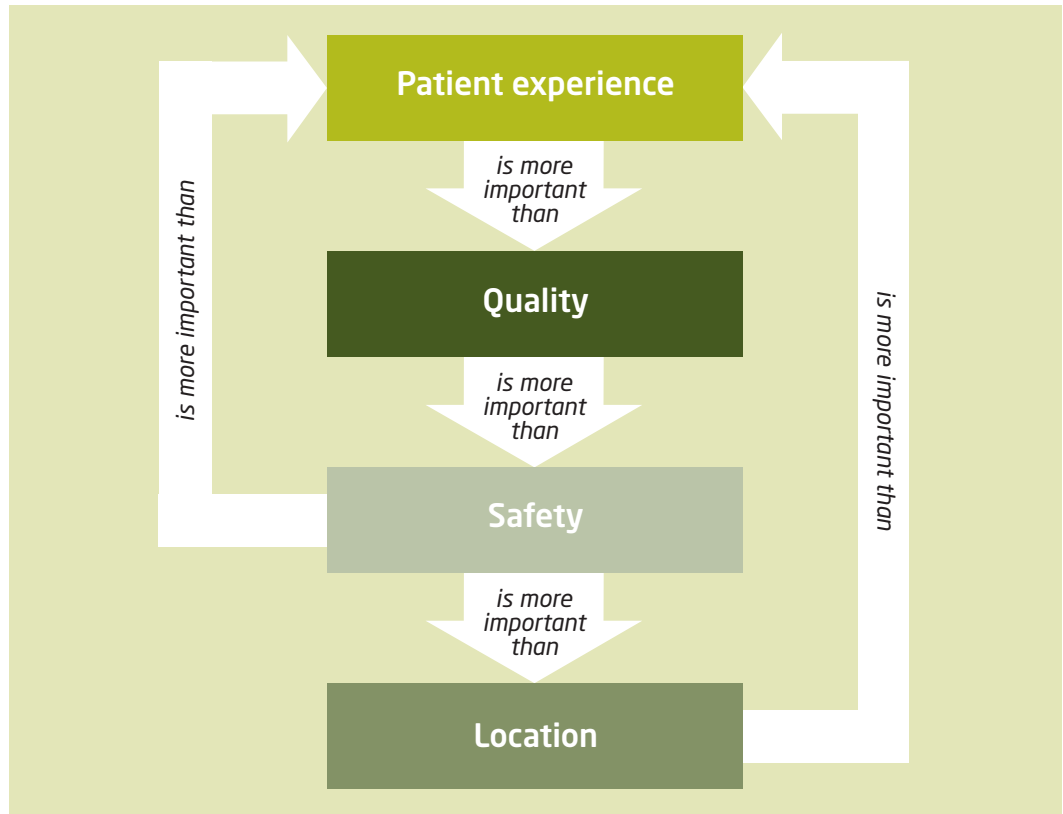
What choices did people make, and who made the 'best' choices?

What choices people made in relation to quality

- Approximately half of participants chose the highest quality hospital across all three choices.

- The vast majority of participants chose the best or second best hospital across all three choices.

- Those with a high level of numeracy were more likely to choose the highest quality hospital.

- Differences in the proportion of participants who chose the highest quality hospital by age, gender and health service use were not significant.

We aimed to find out who chose which hospital when presented with different scorecard designs. We were interested in what proportion of participants chose the 'best' hospital – that is, the hospital with the best performance on clinical quality and safety indicators. Based on our focus group findings, we expected different scorecard designs to have different effects depending on the participants' demographic characteristics. First, we present results for the sample as a whole, then an analysis of differences by demographic characteristics.

Participants chose the highest quality hospital most often in the second hospital choice (C2), where all the information was revealed, and minutes after they had made the first hospital choice (C1). They chose the best hospital least often when making the first hospital choice (C1), which was the first time they were presented with the choice, and when information was hidden.

Approximately half of participants chose the highest quality hospital across all three choices; 47 per cent chose the highest quality hospital in C1, rising to 64 per cent in C2. In phase 2 (C3), when information was hidden, as in C1, 56 per cent chose the best hospital. This percentage is statistically higher than in the corresponding choice task in C1.

Figure 13, overleaf, shows the percentage of participants who chose not just the highest ranked hospital by quality but also hospitals that were ranked second, third, fourth and fifth for quality. Most participants chose the best or second best hospital for all three choices: 85 per cent in C1; 92 per cent in C2; and 87 per cent in C3.

Table 3 (*see* page 47) shows the proportion of participants choosing the highest quality hospital by demographic group and scorecard design, for hospital choice 1 (C1). There were significant differences by numeracy, scorecard design, and likelihood of using the internet to search for health information. Only numeracy and scorecard design remained significant after controlling for other factors in the regression model. Differences in the proportion of participants who chose the highest quality hospital by age, gender and health service use were not significant, nor were these significant predictors of choosing the highest quality hospital. We repeated the analysis using a dichotomous variable for age (under 65 years vs 65 and older), but this did not change the fact that the differences were not significant. Only 7 per cent of the total sample were aged over 65.

Those with a high level of numeracy (who answered two or three questions correctly) were more likely to choose the highest quality hospital – half (50 per cent) did so, compared with 36 per cent of those with a very low level of numeracy (who answered none of the questions correctly). Differences by education were not significant in the

Figure 13 Percentage of participants choosing each hospital (C1, C2 and C3)
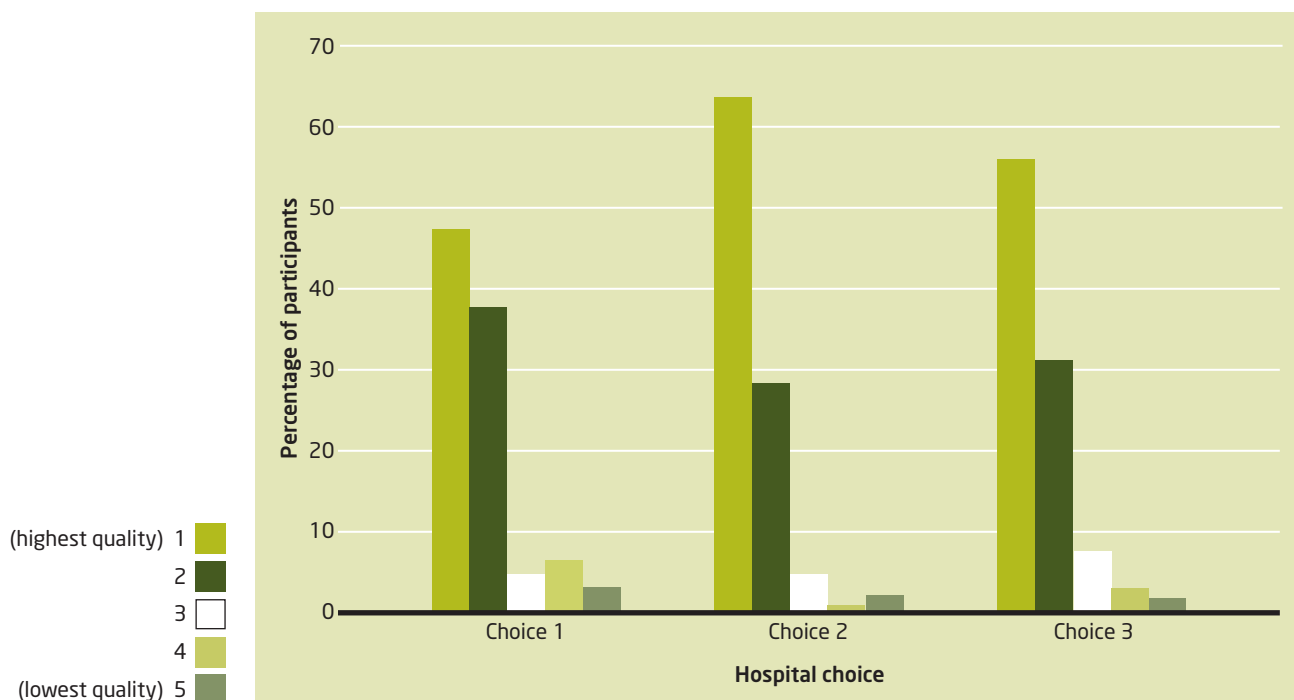
bivariate analysis, and when we included education instead of numeracy in the regression model, it was not significantly associated with choosing the highest quality hospital, and there were no significant differences between those with different levels of education. However, the direction of the responses is opposite to what we expected: those with higher qualifications were less likely to choose the highest quality hospital than those with no or fewer qualifications. Numeracy remained significant in the regression model: those with high levels of numeracy were more likely to choose the highest quality hospital compared with those with very low levels of numeracy (Odds Ratio (OR) = 1.76 and 1.78 respectively for those who answered two or three questions correctly).

Those who said they were very likely to search for health information online were significantly more likely to choose the highest quality hospital than those who said they were very unlikely to do so (50 per cent vs 42 per cent). Nearly half of those who reported being 'always connected' to the internet or 'at least daily' chose the highest quality hospital compared with just 11 per cent of those who rarely accessed the internet (1–3 times per month). It is important to remember that this was an online panel, and therefore the sample is likely to be more internet-savvy than the general population. These were not, however, significant in the overall model after controlling for other factors.

The scorecard design also made a significant difference to whether participants chose the highest quality hospital (*see* below for further analysis). This remained significant after controlling for other variables in the binary logistic regression (*see* Table 4, p 48).

We repeated the analysis for hospital choices 2 and 3 (C2 and C3) (results not shown). Results were similar to those for hospital choice 1 (C1), although the proportion choosing the highest quality hospital was higher across the board (*see* 'Practice, and learning by doing', on page 57). Numeracy, likelihood of searching for health information online, and scorecard design were all significant in the bivariate analysis. Numeracy and scorecard design remained significant in the multivariate analysis, and the direction of the relationships was also stable. Choice of the highest quality hospital by education was again varied, with no obvious direction of association by level of education. However,

those with primary education were significantly more likely to choose the highest quality hospital than those with no education (OR 1.85, p=0.046).

In hospital choice 3 (C3), none of the differences were significant in the bivariate analysis. The oldest age group were marginally less likely to choose the highest quality hospital (p<0.1), and those with a high level of numeracy were slightly more likely to do so in the regression analysis, but no factors reached significance in predicting the overall model.

Of the different variables, numeracy was the best predictor of choosing the highest quality hospital in phase 1. As Figure 14 (*see* page 49) shows, highly numerate participants (who answered all three questions correctly) had the highest proportion of 'best' choices in C1 and C2. Those with very low numeracy (who answered no questions correctly) consistently had the lowest proportion of best choices across the three tasks. This result confirms the crucial importance of numeracy in decision-making in general, and health decisions in particular (Peters *et al* 2006; Hibbard *et al* 2007). Recent estimates suggest that 75 per cent of the adult population of working age have numeracy skills below the level of a good pass at GCSE (House of Commons Public Accounts Committee 2009).

Table 3 Percentage of participants who chose the highest quality hospital in hospital choice 1 (C1) by key demographic variables

Variable	Category	No	%	Significance (p value)
Total		1,168	47	
Gender	Male	529	48	0.593
	Female	639	47	
Age	18–34	491	47	0.577
	35–50	122	47	
	51–65	470	49	
	65+	85	41	
Education	None	106	54	0.225
	Primary school	125	53	
	Secondary/high school	293	43	
	College	193	49	
	Graduate/postgraduate	451	46	
Internet use	Rare to regular	804	47	0.646
	Always connected	362	48	
Numeracy (correct answers)	0	181	36	0.005
	1	277	46	
	2	289	50	
	3	421	51	
Seek health information online	Very unlikely	132	42	0.014
	Very likely	888	50	
GP visits in past year	0	178	50	0.63
	1–3 times	595	48	
	4–6 times	233	45	
	7+ times	160	45	
Hospital visits in past year	0	391	50	0.271
	1–3 times	626	46	
	4–6 times	100	47	
	7+ times	45	36	
Offered hospital choice	Not in the past year	770	48	0.819
	In the past year	394	47	
Scorecard design	Distance sort	138	48	0.000
	Distance sort with information	118	58	
	Quality sort	134	37	
	Quality sort with information	120	34	
	Pre-select nudge	215	45	
	Preference nudge	220	49	
	Feedback nudge	223	56	

Table 4 Binary logistic regression model of choice of highest quality hospital in hospital choice 1 (C1)

Variable	Category	Significance (p value)	Odds ratio	Lower 95% confidence interval (CI)	Upper 95% CI
Gender	(Female)				
	Male	0.769	0.961	0.736	1.254
Age	(18–34)				
	35–50	0.44	1.226	0.731	2.054
	51–65	0.729	1.116	0.599	2.081
	65+	0.212	1.392	0.828	2.338
Internet use	(Rare to regular)				
	Always connected	0.69	1.057	0.805	1.388
Numeracy (correct answers)*	(0)				
	1	0.052	1.541	0.996	2.383
	2*	0.01	1.757	1.142	2.702
	3*	0.007	1.777	1.173	2.692
Seek health information online	(Very unlikely)				
	Very likely	0.087	0.713	0.483	1.051
GP visits in past year	(0)				
	1–3 times	0.287	0.813	0.556	1.189
	4–6 times	0.203	0.743	0.47	1.174
	7+ times	0.512	0.842	0.504	1.407
Hospital visits in past year	(0)				
	1–3 times	0.601	0.925	0.69	1.24
	4–6 times	0.67	0.896	0.542	1.483
	7+ times	0.446	0.757	0.369	1.551
Offered hospital choice	(Not in the past year)				
	In the past year	0.369	0.882	0.671	1.16
Scorecard design*	(Distance sort)				
	Distance sort with information	0.083	1.603	0.941	2.731
	Quality sort	0.281	0.75	0.445	1.265
	Quality sort with information	0.101	0.634	0.368	1.093
	Pre-select nudge	0.579	1.142	0.715	1.825
	Preference nudge	0.398	1.22	0.769	1.935
	Feedback nudge	0.081	1.511	0.95	2.404
Constant		0.183	0.599		

Note: Dependent variable: choice of highest quality hospital. All variables included in the model are shown. The results are expressed as the odds ratio that a particular category of participants chose the highest quality hospital compared with the comparison group (in brackets).

* Significant at 5 per cent level (p<0.05)

Figure 14 Percentage of participants who chose the highest quality hospital in each choice (C1, C2 and C3) by numeracy
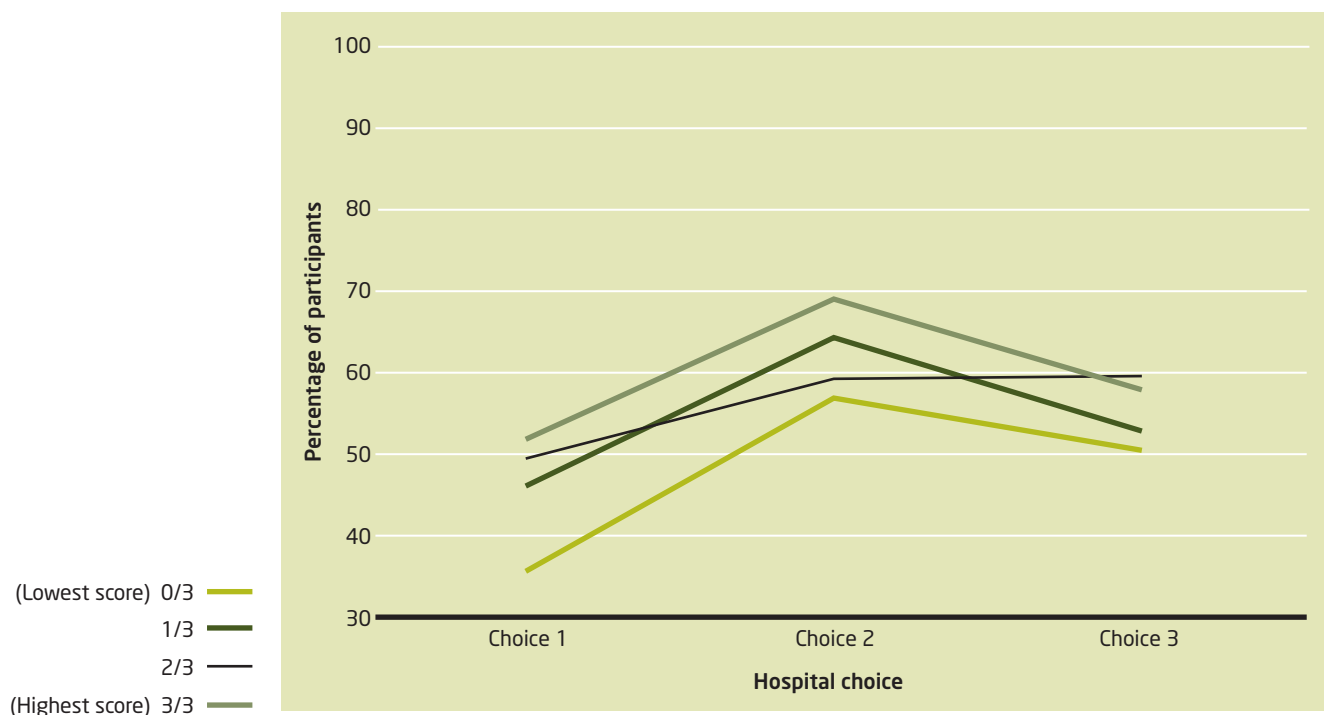
How scorecard design affected quality of choice

How scorecard design affected decision-making

- Participants who had the mindful nudge (preference and feedback) scorecards made the best choices.

- Participants who had the mindless nudge (pre-select) scorecard made the worst choices.

- Some of the presentation formats were more effective in helping certain demographic groups to choose the highest quality hospital.

- Younger participants appear to have been helped by the nudges, particularly those with higher education. For older participants, and younger participants with low education, only the feedback nudge significantly helped.

- Participants who saw hospitals sorted by quality chose significantly worse than those who saw hospitals sorted by distance.

We were interested in how participants' choice of hospital was influenced by the design of the scorecard. We also expected it to have different effects depending on the personal characteristics of the respondent – for example, age, education, numeracy skills, and so on. The effects of scorecard design on quality of choice are presented first for all participants, then broken down by demographic characteristics. As described earlier, 'older' participants are those 51 years of age or over and 'younger' participants are

those 50 years of age and under and 'high education' refers to those with at least some college education, whereas 'low education' refers to those up to secondary education (*see* Methodology p 35).

General effect of design and nudges on highest quality choices

Nudges (pre-select, preference and feedback) marginally improved hospital choices. Those participants who were nudged chose the highest quality hospital marginally more often (50 per cent on average) in C1 compared to participants who had control scorecards (distance or quality sort), 44 per cent of whom chose the highest quality hospital (*see* Table 5 below). Compared to those participants who saw hospitals sorted by quality, those participants who had the preference and feedback nudges were significantly more likely to chose the highest quality hospital in C1 ($p<0.05$). Participants who had the mindless nudge (pre-select) scorecard made the worst choices. They chose only as well as those with control scorecards in C1, and worse than those with control scorecards in C2. Their choices were worse on average than those with mindful nudges in both C1 (53 per cent with mindful nudges vs 45 per cent with pre-select nudge) and C2 (67 per cent with mindful nudges vs 60 per cent with pre-select).

Of the two mindful nudges (preference and feedback), feedback had a more substantial influence on choice of the highest quality hospital. Compared with control scorecards, feedback scorecards led participants to:

- choose the highest quality hospital significantly more often in C1 (56 per cent vs 44 per cent)

- choose the highest quality hospital significantly more often in C2 (71 per cent vs 63 per cent).

Based on this evidence, a pre-select nudge does not appear to be a successful method for improving choices. However, mindful nudges (the feedback nudge in particular) led to better choices; more participants chose the highest quality hospital in choices 1 and 2, both when information was concealed and when it was revealed.

Table 5 Percentage of participants choosing the highest quality hospital in choices 1 and 2 (C1 and C2) by scorecard design

	Choice 1 %	Choice 2 %
Control scorecards		
Distance sort	48	75
Distance sort with information	58	72
Quality sort	37	56
Quality sort with information	34	48
Average across control scorecards	44	63
Nudge scorecards		
Pre-select (mindless)	45	60
Preference (mindful)	49	62
Feedback (mindful)	56	71
Average across mindful nudges	53	67
Average across nudge scorecards	50	64

Effect of scorecard design and nudges on demographic groups

As Table 6, below, shows, the scorecard design had different effects on different demographic groups.

While the majority of older participants (aged 51 or above) who had distance sort scorecards for hospital choice 1 (C1) chose the highest quality hospital, regardless of level of education (60 per cent and 59 per cent), the younger participants (aged 50 or below), particularly those with higher levels of education, were less likely to choose the highest quality hospital (35 per cent and 47 per cent). As previously noted, overall, participants were less likely to choose the highest quality hospital with quality sort. However, the pattern of differences between young and old was similar, with older participants more likely to choose the highest quality hospital than younger ones (though these differences were not statistically significant).

The differential impact of these scorecard designs was not apparent in hospital choice 2 (C2). There could be a number of explanations as to why younger participants did better in C2 than C1. First, they are less likely than older participants to have had previous experience of hospital care, and therefore may have been less familiar with the types of factors presented for consideration. They may, therefore, have learned more from the first exercise. Second, they may have found the task much easier when the information was revealed. However, it would be surprising to find that the blurred information, and the requirement to mouse click to reveal it, was less of a barrier for older than younger participants. Finally, younger participants may have given weight to different factors more systematically than older participants. Although the values used in the scorecards in C1

Table 6 Percentage of participants choosing the highest quality hospital by scorecard design and demographic characteristics in hospital choices 1 and 2 (C1 and C2)

	Older – high education %	Older – low education %	Younger – high education %	Younger – low education %	Average %	Significance (p value)
Choice 1						
Distance sort	60	59	35	47	47.8	**0.093**
Distance sort with information	67	62	60	45	58	0.357
Quality sort	47	44	28	32	37	0.295
Quality sort with information	43	28	33	33	34	0.678
Pre-select nudge	38	48	50	46	45	0.565
Preference nudge	37	42	54	60	49	**0.044**
Feedback nudge	57	55	59	52	56	0.887
Average	48	48	46	47	47	
p value	**0.026**	0.127	**0.007**	0.142		
Choice 2						
Distance sort	70	74	73	80	74	0.834
Distance sort with information	67	69	77	76	72	0.778
Quality sort	65	59	46	56	56	0.441
Quality sort with information	61	45	44	41	48	0.445
Pre-select nudge	51	67	70	56	60	0.124
Preference nudge	66	68	63	52	62	0.274
Feedback nudge	77	73	69	65	71	0.511
Average	65	66	64	60	63	
p value	**0.088**	0.237	**0.011**	**0.019**		

Note: Chi square tests were conducted for each column (ie, among those in a particular age/education grouping, were the differences by scorecard significant?) and row (ie, among those who had a particular scorecard, were the differences by age/education groupings significant?). Figures highlighted in bold are significant at less than 10 per cent level (p<0.1). The numbers in each cell were small, and therefore it was difficult for differences to achieve statistical significance.

and C2 were different, the second ranked hospital did better on patient experience than the hospital ranked first in both scorecards, so if younger participants gave more weight to the patient experience, it would not have made a difference between C1 and C2.

Among the nudges, the differential impact of the preference nudge was significant. The proportion of older participants who chose the highest quality hospital when nudged in this way (37 per cent of those with high education and 42 per cent with low education) was less than among the younger ones (54 per cent of those with high education and 60 per cent with low education). Among both younger and older participants, the higher educated were less likely than the lower educated to be helped by this nudge. The other nudges had no significant observable differential impacts. However, as previously noted, the feedback nudge helped a greater proportion of participants choose the highest quality hospital than the pre-select nudge.

Some of the presentation formats were more effective in helping certain demographic groups to choose the highest quality hospital (*see* Table 6 on page 51).

Among older participants with high levels of education, 48 per cent (on average) chose the highest quality hospital in C1, rising to 65 per cent in C2. Differences by scorecard design were significant in both C1 and C2. A higher proportion of those who had the distance sort and distance sort with information scorecards chose the highest quality hospital compared with those who had quality sort and quality sort with information scorecards. Those who had the feedback nudge did better than average (57 per cent), whereas those who had the pre-select and preference nudges did worse than average (38 per cent and 37 per cent respectively). These differences were also apparent in C2, except that the proportion of those who had the preference nudge in C2 was no worse than average (66 per cent).

There were no significant differences among older participants with low levels of education in terms of the proportion who chose the highest quality hospital by scorecard design in either C1 or C2. However, the pattern of differences was very similar to that for older participants with high levels of education.

Among younger participants with high levels of education, the differences by scorecard were also significant in both C1 and C2. On average, 46 per cent of this group chose the highest quality hospital in C1 and 64 per cent in C2. The nudges appeared to help this group, as did the distance sort with information; 60 per cent of those who had this scorecard with extra information to guide them about what factors are important chose the highest quality hospital, compared with just 35 per cent of those who had the distance sort (that is, they simply saw the scorecard without first being given additional information). This group appeared to fare particularly badly with the quality sort and quality sort with information scorecards. Similar differences were apparent in C2, except that the proportion of those using distance sort in C2 that chose the highest quality hospital was higher than average (73 per cent).

There were no significant differences among younger participants with low education in the proportion who chose the highest quality hospital by scorecard design in C1, but the differences were significant in C2. On average, 47 per cent of this group chose the highest quality hospital in C1, and 60 per cent in C2. In contrast to the higher educated, those who had distance sort did no worse than average (47 per cent) in C1. For this group, only the feedback nudge appeared to help; 65 per cent of participants who had this nudge chose the highest quality hospital.

In summary, some scorecard designs made it more likely that particular groups would choose the highest quality hospital, but also made it less likely that others would choose the highest quality hospital. In general, younger participants appear to have been helped by the nudges, particularly those with higher education. For older participants, and younger participants with low education, only the feedback nudge significantly helped.

Older respondents' familiarity with hospital care and the factors used to measure quality of care might explain why more of them chose the highest quality hospital than the younger participants in hospital choice 1 (C1) when presented with a standard scorecard sorted by distance. However, younger participants appear to have learned from the first exercise, and did as well as older participants in the second choice (C2). In C1, the younger participants who were informed about differences in the quality of hospital care and were asked to think about and indicate their preferences did better than older participants who had the same nudge. The fact that younger participants with high levels of education who got similar information about differences in quality of hospital care were also more likely to choose the highest quality provider suggests that they benefited from being made aware of quality differences between hospitals, and being made to think about what might be important when selecting a hospital for treatment. It is likely that older participants may have thought about these issues previously and be more aware of such differences.

Prioritising quality in scorecards

The impact of sorting hospitals by quality had a surprising effect on the quality of choices. Participants who saw hospitals sorted by quality chose significantly *worse*, both in C1 (36 per cent vs 53 per cent with distance sort) and C2 (52 per cent vs 73 per cent with distance sort). If the scorecard is presented with one hospital per column and in a horizontal series (as we did in our experiment, to reflect NHS Choices), then sorting hospitals by quality – although it might *appear* to be most helpful – will most likely not lead to better choices. Instead of sorting by *quality*, sorting by *distance* improves hospital choices.

Research suggests two explanations for this counter-intuitive result. Research on decision behaviour in online environments finds that people over-search when search 'cost' is lowered by screening tools (Diehl and Zauberman 2005). Because people tend to over-search, a screening tool which first presents the highest quality options (as in the case of quality sort) will lead people to make worse choices than a screening tool which presents the highest quality option later on (as was the case for the distance sort scorecard). In addition, research on spatial perception suggests that when options (in our experiment, hospitals) are presented simultaneously and horizontally, people have an unconscious attraction towards the centre of the screen (Fitzsimons *et al* 2002; Valenzuela and Raghubir 2010). The attraction to the centre claims not to be driven by attention (our heatmaps show information presented first receives most attention), but rather, it occurs because it might seem to 'simplify' trade-offs for people (ie, options presented in the middle offer 'balance' between 'extreme' options at the edges of the screen). This centrality effect could have led participants to choose the highest quality hospital more often from the distance sort scorecard (where the highest quality hospital was close to the middle, in fourth position) than from the quality sort scorecard (where it appeared first, and therefore had an 'extreme' positioning).

How satisfied and confident did people feel?

How satisfied were people with their decision

■ While mindful nudges improve the choices made by those in the young, highly educated group, at the same time they also decrease their satisfaction and feeling of being informed.

Participants were asked to rate how satisfied they were with the choices they made on a scale from 1 to 10. We used responses received from participants after they made hospital choice 1 (C1) to understand how they felt about the choice they made and, in particular, how satisfied they were with their choice. Findings are presented first by overall sample characteristics, and then by demographic group and scorecard design.

General effect of demographics

Table 7, opposite, shows the proportion of participants who reported being satisfied with their choice (hospital choice 1) by demographic group. There were significant differences by level of education and numeracy; those with no educational qualifications were less satisfied than those with some education, but interestingly, those with some college or graduate/postgraduate education were less satisfied than those with only primary school or secondary/high school education. Those who answered two or three of the numeracy questions correctly were more satisfied than those who answered only one or two correctly.

Differences by age, gender, and internet use were not significant. However, those who said they would use the internet to search for health information were more satisfied than those who would not (79 per cent vs 72 per cent). While differences by number of GP and hospital visits in the past year were not significant, people who had been offered a choice of hospital in the past year were significantly more satisfied than those who had not (81 per cent vs 76 per cent). Men were more likely to be satisfied than women (80 per cent vs 76 per cent), though this was only marginally significant (p=0.10).

After controlling for other factors, numeracy, scorecard design, and whether someone had been offered a choice in the past year were significantly associated with satisfaction (*see* Table 8, p 56). We explore the impact of scorecard design on satisfaction in the following section.

Participants who reported having been offered a choice of hospital previously were nearly 50 per cent more likely to be very satisfied with their choice than those who had never been offered a choice. This suggests that those who have experience of making a similar decision in the past found the task more satisfying, which is consistent with our findings on learning (*see* 'Practice, and learning by doing' on page 57).

Looking at the variable 'hospital visits in the past year', those who had been referred to hospital seven times or more in the past year (or had a family member in this category) were approximately half as likely to be very satisfied as those who had not been referred to a hospital in the past year.

Those with higher levels of numeracy (answering two or three questions correctly) were around twice as likely to be very satisfied as those who had low levels of numeracy (ie, those who answered no or only one question correctly). Education was significant when this was included in the model instead of numeracy. Those with secondary/high

school education were more than three times as likely to say they were very satisfied as those with no education (OR=3.17), those with primary school education were more than twice as likely (OR=2.30), and those with graduate/postgraduate education were nearly twice as likely (OR=1.68).

Analysis was repeated for hospital choices 2 and 3 (results not shown). Overall, satisfaction with the choice declined between choice 1 and 2, and then stabilised. Differences by numeracy remained significant in the bivariate analysis. Differences in satisfaction by scorecard design, education, and experience of choosing a hospital were no longer significant in choice 2. There were no significant differences in satisfaction in choice 3. As we will see below, there was a learning effect; it may be that those groups who found the choice challenging were dissatisfied with their initial experience of choosing a hospital, but as they became more familiar with the task, these differences disappeared. However, making people more aware of differences between providers (either by using mindful nudges or by putting them through a choice task) appears to reduce satisfaction with the choice made.

Table 7 Percentage of participants who were very satisfied (8 and above) in hospital choice 1 (C1) by key demographic variables

Variable	Category	No	%	Significance (p value)
Total		1,135	37	
Gender	Male	516	80	0.104
	Female	619	76	
Age	18–34	480	75	0.267
	35–50	121	79	
	51–65	452	79	
	65+	82	80	
Education	None	130	66	0.005
	Primary school	119	81	
	Secondary/high school	282	83	
	College	188	74	
	Graduate/postgraduate	443	77	
Internet use	Rare to regular	780	78	0.619
	Always connected	353	76	
Numeracy (correct answers)	0	171	68	0.000
	1	266	72	
	2	284	83	
	3	414	81	
Seek health information online	Very unlikely	124	72	0.066
	Very likely	871	79	
GP visits in past year	0	173	77	0.923
	1–3 times	577	78	
	4–6 times	228	76	
	7+ times	155	78	
Hospital visits in past year	0	383	78	0.125
	1–3 times	606	78	
	4–6 times	97	74	
	7+ times	44	64	
Offered hospital choice	Not in the past year	744	76	0.047
	In the past year	387	81	
Scorecard design	Distance sort	135	77	0.009
	Distance sort with information	117	86	
	Quality sort	133	82	
	Quality sort with information	118	79	
	Pre-select nudge	210	75	
	Preference nudge	214	69	
	Feedback nudge	208	80	

Table 8 Binary logistic regression model of satisfaction with hospital choice (C1)

Variable	Category	Significance (p value)	Odds ratio	Lower 95% confidence interval (CI)	Upper 95% CI
Gender	(Female)				
	Male	0.745	1.056	0.759	1.47
Age	(18–34)				
	35–50	0.664	0.867	0.455	1.652
	51–65	0.804	1.104	0.506	2.412
	65+	0.75	1.112	0.579	2.135
Internet use	(Rare to regular)				
	Always connected	0.914	1.019	0.729	1.423
Numeracy (correct answers)*	(0)				
	1	0.994	0.998	0.615	1.62
	2*	0.004	2.157	1.287	3.614
	3*	0.028	1.713	1.06	2.766
Seek health information online	(Very unlikely)				
	Very likely	0.133	0.711	0.455	1.11
GP visits in past year	(0)				
	1–3 times	0.94	1.018	0.639	1.623
	4–6 times	0.818	1.068	0.612	1.863
	7+ times	0.306	1.394	0.739	2.63
Hospital visits in past year	(0)				
	1–3 times	0.452	0.869	0.603	1.253
	4–6 times	0.17	0.658	0.362	1.197
	7+ times*	0.009	0.36	0.167	0.774
Offered hospital choice*	(Not in the past year)				
	In the past year*	0.038	1.442	1.02	2.038
Scorecard design*	(Distance sort)				
	Distance sort with information	0.061	1.989	0.97	4.079
	Quality sort	0.149	1.629	0.84	3.159
	Quality sort with information	0.474	1.269	0.661	2.436
	Pre-select nudge	0.731	0.907	0.519	1.586
	Preference nudge	0.388	0.787	0.457	1.356
	Feedback nudge	0.431	1.26	0.709	2.238
Constant		0.091	2.204		

Note: Dependent variable: satisfaction with choice. All variables included in the model are shown. The results are expressed as the odds ratio that a particular category of respondents were very satisfied with their choice compared with the comparison group (in brackets).

* Significant at 5 per cent level

Effect of demographics and scorecard design

Our results show greater dissatisfaction with their hospital choices among participants who were nudged compared with those who received control scorecards. This is mainly the case in the young, highly educated group – the same group that is helped most by nudges in choosing the highest quality hospital (*see* Table 9, opposite). Compared to those using control scorecards, young, highly educated participants who were nudged were:

- significantly[10] less satisfied in hospital choice 1 (7.70 vs 8.21)

- significantly less informed in hospital choice 1 (7.55 vs 8.08)

- significantly less satisfied in hospital choice 2 (7.69 vs 8.27)

- significantly less informed in hospital choice 2 (7.56 vs 8.17).

While mindful nudges improve the choices made by those in the young, highly educated group, at the same time, they also decrease their satisfaction and feeling of being

[10] statistically different at p<0.5 (unpaired t-tests)

Table 9 Percentage of participants who were very satisfied (8 and above) by scorecard design and demographic characteristics for hospital choices 1 and 2 (C1 and C2)

Choice 1	Older: high education %	Older: low education %	Younger: high education %	Younger: low education %	Average %	Significance (p value)
Distance sort	67	81	73	83	75	0.390
Distance sort with information	80	90	90	83	86	0.613
Quality sort	74	81	87	82	81	0.519
Quality sort with information	89	79	78	63	78	0.135
Pre-select nudge	79	76	66	69	73	0.380
Preference nudge	68	71	63	68	67	0.868
Feedback nudge	77	73	72	77	75	0.907
Average	76	78	74	74	75	
p value	0.303	0.601	**0.043**	0.291		
Choice 2						
Distance sort	67	67	61	80	67	0.362
Distance sort with information	60	66	80	66	68	0.389
Quality sort	59	56	82	79	70	**0.031**
Quality sort with information	75	69	64	56	66	0.475
Pre-select nudge	63	76	70	62	67	0.403
Preference nudge	65	61	58	58	60	0.864
Feedback nudge	68	75	59	50	62	**0.053**
Average	65	68	66	62	65	
p value	0.873	0.552	**0.082**	**0.040**		

Note: Chi square tests were conducted for each column (ie, among those in a particular age/education grouping, were the differences by scorecard significant?) and row (ie, among those who had a particular scorecard, were the differences by age/education groupings significant?). Figures highlighted in bold are significant at less than 10 per cent level (p<0.1). The numbers in each cell were small, and therefore it was difficult for differences to achieve statistical significance.

informed. This is most likely an effect of the additional thinking and introspection that mindful nudges, and preference nudges in particular, require. In the context of difficult choices, this extra thinking and introspection generates doubts and dissatisfaction, as has long been shown in the psychological literature (Wilson and Schooler 1991; Wilson *et al* 1993).

Practice, and learning by doing

Practice

- Learning occurred across all scorecard designs. There were no significant differences across the nudge scorecards, suggesting that learning occurred equally often with different nudges.

- Learning occurred equally often regardless of age, education or numeracy.

- The effects of nudges on choice appear to be short-lived, suggesting nudges need to be used repeatedly.

- Mindful nudges had a lasting effect on increasing people's awareness even when the nudges were removed.

- The long-lasting effect of mindful nudges is a feeling of greater confidence in the ability to make the choice, and greater awareness that information is available.

We sought to understand whether the ability of participants in both phase 1 and phase 2 to make a high-quality choice improved over time, and the role of repeated practice. As shown in Figure 14 (*see* page 49), across all scorecard designs, the proportion of participants selecting the highest quality hospital significantly improved between hospital choice 1 (C1) and hospital choice 2 (C2), and we infer from this that some sort of learning took place.

We analysed the differences between C1 and C2, and any improvement that emerged in phase 1 where nudges were and were not present. To have a richer picture of the improvement, we examined how choice improved across the whole range of hospitals chosen. Consequently, learning was operationalised as the difference in average ranking from C1 to C2, where the maximum learning is +4 (from lowest ranking in C1 to the highest ranking in C2) and minimum learning is –4 (from highest ranking in C1 to the lowest ranking in C2), and the lower the rank, the better.[11]

Learning occurred across all scorecard designs. There were no significant differences across the nudge scorecards, suggesting that learning occurred equally often with different nudges. Marginally, the most significant learning between C1 and C2 was found in the distance sort scorecard when compared to quality sort (this is most likely due to the very poor performance of participants in quality sort). Learning in the controls scorecard is an important result, as this means that learning is not nudge-specific, and occurs in each choice scorecard. In addition, there were no notable differences in learning by demographic variable – so learning occurred equally often regardless of age, education or numeracy.

The learning we observed in phase 1 could therefore be due to two possible explanations: repeated practice, or the fact that C2 was 'easier' because the information was all revealed (information did not have to be held in the person's memory, as a similar task had just been completed). To disentangle these two points, we turn to phase 2, where the choice of hospital was offered after an interval of between two and four weeks. Information about the hospital was hidden in the same way as in choice 1.

Comparing all scorecards in phase 1 with all scorecards in phase 2, we find that:

■ participants in phase 2 (N=740) chose the highest quality hospital more often in phase 2 (56 per cent) than in phase 1 (49 per cent)

■ the average rank of the chosen hospital is lower (=higher quality) in phase 2 (rank=1.63) than in phase 1 (rank=1.76).

We sought to understand whether there were any long-lasting effects of the scorecard designs that participants were exposed to in phase 1. The aim was to see if the design and the particular nudge used had any effect in phase 2, where no nudge was provided. Table 10, opposite, shows in order of 'most learning' the difference between the two comparable choice tasks in phase 1 and phase 2, between hospital choice 1 (C1) and hospital choice 3 (C3) (both hidden information).

Participants who had control scorecards (who had no nudges in phase 1 and performed worse on average) improved more than those who had nudge scorecards in phase 1 (*see* Table 11, opposite). This suggests that the 'nudge effect' on improving choice outcome is short-lived, and requires a nudge to be omnipresent. Choice is extremely context-dependent, and in online settings, is powerfully shaped by the format of the information presented.

[11] Rank of '1' means the best hospital was chosen; rank of '5' means the worst hospital was chosen.

Table 10 Difference in average rank of hospital chosen between hospital choices 1 and 3 (C1 and C3)

Scorecard design	Learning between C1 and C3
Feedback nudge[12]	– 0.06
Distance sort	+ 0.03
Preference nudge	+ 0.06
Pre-select nudge	+ 0.21
Quality sort	+ 0.26

Table 11 Percentage of participants choosing the highest quality hospital in choices 1, 2 and 3 (C1, C2 and C3) conditional on phase 1 scorecard design

	C1 %	C2 %	C3 %
Distance sort	48	75	56
Distance sort with information	58	72	52
Quality sort	37	56	61
Quality sort with information	34	48	61
Pre-select nudge	45	60	64
Preference nudge	49	62	49
Feedback nudge	56	71	52

The mindless pre-select nudge exerts a long-lasting effect on quality of choice outcome compared with mindful nudges. Participants who had the mindless pre-select nudge scorecard in phase 1 chose in phase 2 the highest quality hospital significantly more often (64 per cent) than those who had mindful nudges in phase 1 (50 per cent). This suggests that the more cognitively challenging nudge, rather than enhancing people's ability to choose the highest quality provider again in future, actually diminishes it. Participants who had the pre-select nudge in phase 1 made better choices when not given any help (in the form of a nudge) in phase 2, whereas those who had been made to think more about their preferences in phase 1 appeared to make worse decisions when not given any nudge in phase 2. Contrary to what we expected, mindful nudges do not appear to improve people's ability to choose higher quality hospitals; however, they may lead people to make more 'informed' choices – that is, in line with their own preferences. We have not been able to test this hypothesis here though.

Long-lasting effects on increasing people's awareness

Nudges, and, in particular, mindful nudges, had a lasting effect on increasing people's awareness even when the nudges were removed. Participants who had mindful nudges in phase 1 felt more confident and informed when making their choices in phase 2. In particular:

- participants who had the feedback nudge in phase 1 felt significantly more confident in phase 2 than those who had control scorecards in phase 1.

[12] Participants in phase 2 were not given any nudge scorecards. Learning reported here is learning that occurred between two phases without nudges being used. That is, the results show how, for example, people who had the feedback scorecard in phase 1 did in phase 2, compared with their performance in phase 1.

Further, participants who had mindful nudges in phase 1 felt marginally more confident than participants who had mindless nudge scorecards that they could:

- get the facts about the hospital choices available to them

- get the facts about the benefits of each hospital

- get the facts about the benefits and disadvantages of each hospital.

Participants who had scorecards with distance sort as opposed to quality sort:

- felt marginally more confident about their decision and more clear about the benefits of the hospital that matter most to them

- were less likely to want to leave the decision to someone else

- were marginally less likely to want to be prompted about the highest quality hospital.

Overall, 37 per cent of participants said they would want a recommendation in future, while more than half (56 per cent) said they would trust a recommendation. Those participants who had control scorecards would trust pre-selection significantly more than participants who had the feedback nudge in phase 1 (ie, they agree more with the statement: 'If a hospital were pre-selected to me as highest quality, I would trust this recommendation'). Those who experienced pre-selection in phase 1 were more likely than average to say they would like a recommendation, but not significantly more likely to trust a recommendation than others (*see* Figure 15 below).

Participants who had mindful nudges in phase 1 made better choices than those who had control scorecards or the pre-select nudge, but this improved choice effect did not last when the nudge was removed. However, the long-lasting effect of mindful nudges

Figure 15 Percentage of participants wanting or trusting a recommendation by scorecard design in phase 1
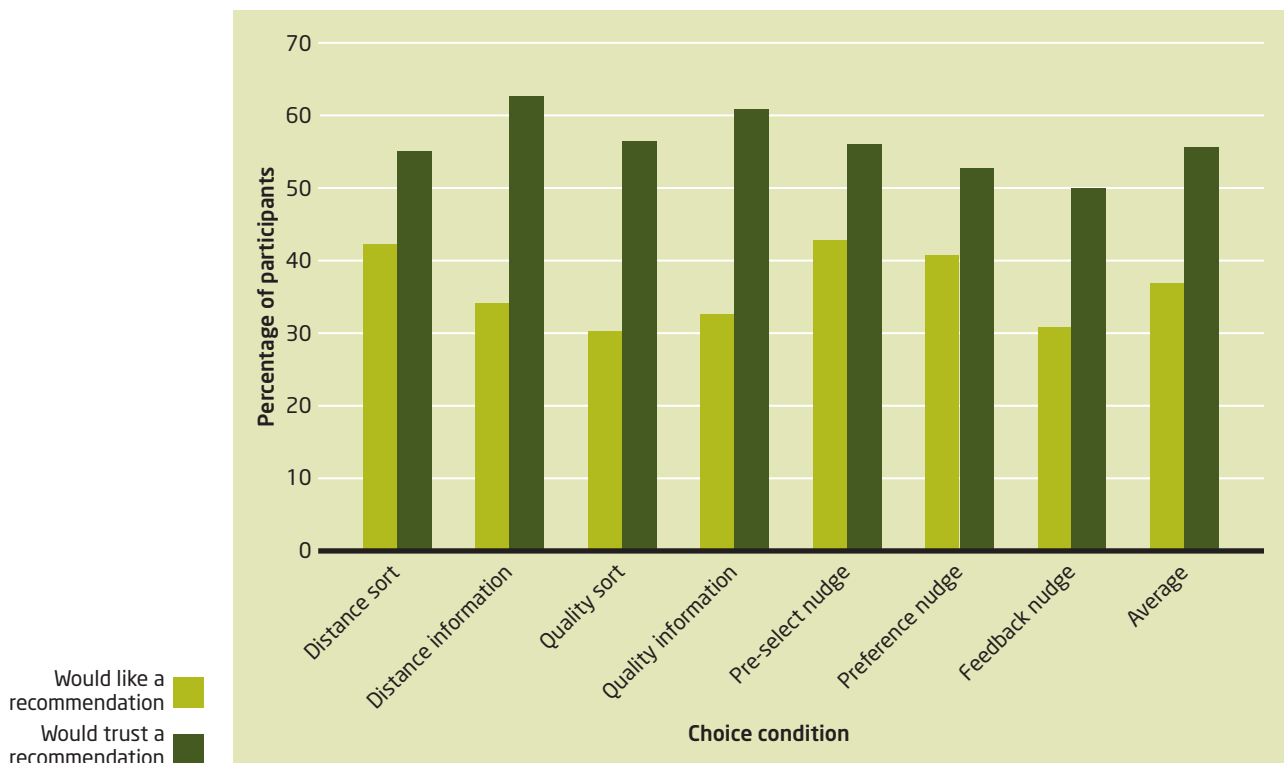
is a feeling of greater confidence in the ability to make the choice, and greater awareness that information is available. There are some long-lasting effects of mindful nudges even after they are removed; these are not reflected in the quality of choice, but rather, in people's attitudes.

Summary and discussion

The findings of our online experiment underline the importance of how information and options are ordered; people spent most time looking at the information at the top of the scorecard and in the first column. It also appears that people who are made to think about their choice and what is important to them (preference nudge) are more systematic in the way they view information and more efficient in their information search. When people are not prompted to think about what is important to them, they spend more time looking at information about particular hospitals, rather than comparing hospitals on specific indicators. We also found that these people paid more systematic attention to quality and safety indicators.

When people are given information about hospital choice and differences between hospitals, they appear to value information about the clinical quality and safety of services more than non-clinical factors. Those who were prompted to think about their choice (preference nudge) were less likely to say they chose the hospital because it was closer to them than those who had a scorecard with hospitals sorted simply by quality or distance. People do not appear to be seeking to maximise quality. They perceive the gap between a poor and a good hospital to be greater than that between a good and an excellent hospital. Participants were therefore not necessarily seeking a hospital that was excellent on indicators of importance, but one that was not poor. This is consistent with findings of other research (Dixon *et al* 2010), that the main reason why patients would choose a non-local hospital is because of a bad experience – in other words, to avoid a poor-performing hospital.

The online experiment showed that making informed choices using comparative information on the quality of hospitals is difficult, even for people who are used to using the internet. The task of processing complex information is easier for those with high levels of numeracy – something that previous research has shown to be important when making health care decisions (Peters *et al* 2007a, b).

While scorecard format (the layout) and design made a difference to people's decisions, the impact was not always as expected, and people responded to the different scorecard formats and nudges in different ways. It is therefore essential that scorecard designs are tested and that different visual formats are available for different groups of people.

Simply setting a default hospital did not improve the choices people made, and was even detrimental for older, highly educated participants. However, mindful nudges, which gave participants feedback on their choice, did increase the proportion of participants who chose the highest quality hospital. This was particularly the case for younger individuals, though this was accompanied by greater dissatisfaction in the younger, highly educated group.

Surprisingly, sorting hospitals by quality did not help people to make better decisions. In fact, those who had scorecards sorted by distance made better decisions. It seems that the local hospital was assessed first (because of the ordering effect described above). This hospital acted as an anchor from which participants were then able to compare the attributes of other hospitals and evaluate whether it was worth going further for something better. We suspect that when the highest quality hospital was presented first,

other hospitals were compared with this, and that other attributes (perhaps including whether a hospital was nearer and performed nearly as well on quality) were used to make a different selection.

Those with high levels of numeracy not only made better choices but were also more likely to be very satisfied with the choice they made. Education, while not associated with quality of choice, *was* associated with level of satisfaction, though the relationship was not linear. Those with secondary or high school education were most likely to be satisfied with their choice. Interestingly, age and frequency of internet use were not predictors of better choices, suggesting that choosing a hospital, while complex, is not necessarily more difficult for older people. However, due to the sampling, participants were likely to be more digitally literate than the average person. The few participants who used the internet rarely were less likely to make the highest quality choice within the sample. Making an informed choice of hospital was not any easier for those who had more frequent contact with the health service, though people who had been offered a choice of hospital were more satisfied with their choice, suggesting that prior experience may play a role in determining how people feel about making a choice of hospital.

There was strong evidence that repeated practice improved participants' ability to make a high-quality choice. Learning occurred across all the scorecard designs. Learning was greater when the choices were made close in time, but overall, there was a longer-term learning effect. We had expected that participants who were prompted to think about their choice in phase 1 (those who had mindful nudge scorecards) would make better choices in phase 2, reflecting a lasting effect on people's ability to choose. In fact, these people made worse choices in phase 2 than those who had other scorecards. It is important to note that none of the nudges were used in phase 2. This suggests that the more cognitively challenging nudge, rather than enhancing patients' ability to choose a higher quality hospital again in future, instead enhances their awareness and confidence about being informed. It would be interesting for future analysis and research to establish the extent to which these nudges help people to choose according to their own preferences.

5 Discussion and conclusions

Taking the findings of the focus groups and online experiments together, we can identify a number of implications for information providers and policy-makers. These should, however, be viewed in the light of a number of limitations of the research, which we briefly discuss below.

Limitations

While we tried to ensure that we had equal samples of young and old participants and those with high and low levels of education in the online experiment, the sample was recruited through an online panel where the incentive for completion was to collect points for shopping online. This means that our participants are likely to be more internet-savvy than the average person. Other research has suggested that it is access rather than skill that is a barrier to use of the internet by older people (Murray *et al* 2003; Stevenson *et al* 2007). Therefore, although our sample may currently have better access to the internet, there is no evidence that they would make systematically different use of the internet for searching and finding health information. In the focus groups, we were able to include a range of participants, including people who were not necessarily used to using the internet. Therefore, overall, our research captures and reflects a diversity of views.

In analysing the results of our online experiment, we pre-selected the highest quality hospital based on our assessment of the quality and safety scores we had assigned to each hospital in the scorecard. This required us to take a normative approach to choice. It is possible to argue that the highest quality choice for each individual is the one that is most closely aligned to their own preferences. For example, if the most important factor for an individual was proximity, then choosing the closest hospital might be the highest quality choice for them, rather than the hospital with the highest scores on quality and safety.

While we collected some data on preferences, these were found to be inconsistent, and it was not possible to clearly establish a set of preferences and match these against choices made. Our intention was to get people to pay more attention to quality and safety and, therefore, for the purposes of this research, we felt it appropriate to use our judgement of the highest quality hospital as the outcome. Given the significant proportion of participants who also chose the second-ranked hospitals, it may be that a number of participants were making different trade-offs, rather than maximising quality and safety, as we did. This needs to be borne in mind when judging the quality of decision-making promoted by the nudges employed in this study.

The research was conducted under experimental conditions, and the choices we were asking people to make were necessarily hypothetical. Participants may have approached the task as a puzzle that they had to solve and get the 'right' answer to, whereas in reality, the task might be approached in a less structured way and other factors might have a greater influence on people's decision-making, which is something we cannot capture in

the experiment. Given that the task was hypothetical, responses about satisfaction do not reflect whether participants were satisfied with the hospital chosen (as they would in a real-life situation), but rather the experience of choosing.

While the findings have strong internal validity, it is difficult to be confident that similar results would be found in a real-life situation. People may take more time to make a real-life decision or, alternatively, may be distracted or 'give up' if the information is too complex to understand or interpret. People may have made a quick decision in order to complete the task, rather than reflect on what is important to them and make a considered view. Participants may have been more susceptible to nudges in the experiment, whereas in real life, people may be less trusting of devices such as pre-selection. However, compared with other experimental studies in decision research, this study recruited UK citizens who know and use the NHS, and therefore has higher external validity than a laboratory experiment with undergraduate students. It was conducted under experimental conditions in order to disentangle the role played by the design we were manipulating. Future research should evaluate the effect of these designs with real patients making real choices.

The online experimental scorecard was designed to be as real as possible, and most of the factors included appear to have been rated similarly to how people report making choices in real life. However, we limited the number of indicators, and waiting times were not included. Previous research suggests that where waiting times vary, they remain a key determinant of choice. The exclusion of this variable might have influenced the choices people made.

Practical implications for policy-makers and information providers

The 2010 White Paper, *Equity and Excellence*, signals that, in future, the government will encourage a range of third parties to provide information to support patient choice, in addition to NHS Choices (Department of Health 2010a, Section 2.14). While we used NHS Choices as the model for the online experiment, our findings should be of interest to a range of organisations (both public and private) that publish online information, including, for example, the Care Quality Commission and Dr Foster. The findings, presented below, are organised into five areas: type of information, presentation, use of numerical data, ordering, and use of nudges.

Type of information

Our research supports the findings of previous research – that people can easily be overwhelmed by the quantity of information provided; less is more when it comes to providing comparative information on the quality of providers (Peters *et al* 2007a, b).

It appears that people will spend only a comparatively short amount of time (a few minutes in many cases) consulting this sort of information online. The heatmaps show that people look systematically at a relatively small proportion of the information shown (in this case, five hospitals and nine indicators). This suggests that ordering is extremely important (*see* below). Tools that eliminate or demote indicators on which there are no (or only negligible) differences between hospitals could reduce the amount of information people need to consider. The web also provides opportunities to tailor information displays to those aspects of care that people indicate are most important to them, thus personalising the type and amount of information presented.

Among focus group participants, older people and those with lower levels of education were keen to see summative measures to help them 'sum up' the quality of each hospital.

However, they were also confused as to the meaning of some of the summative measures provided, such as 'overall quality', when these were presented alongside other specific indicators. Given this confusion, we chose not to include summative measures in our online experiment. Further testing is needed to establish how to combine summative and disaggregated indicators effectively for different groups of people.

Focus group participants were suspicious of 'missing data' or 'data not available'. If data are not available for a particular chosen hospital or indicator, this needs to be framed carefully so that it does not count against a hospital. Other research on hypothetical choices with recently referred patients found that 'missing data' is assumed to indicate poor performance against those indicators (Dixon *et al* 2010). Information providers need to improve the presentation of missing data, providing an explanation of what 'data not available' means, and why they are not available. It should also be an impetus for providers to ensure that they are able to report data wherever possible. The problem of small numbers may be an increasingly common one if information is presented at consultant level, and may result in even more use of 'missing data'. Ideally, 'data not submitted' should be distinguished from 'data not shown'.

Focus group participants also made it clear that they wanted information that was relevant to their scorecard or treatment. They wanted to know the risks they faced by attending a particular hospital, suggesting the need for more tailored decision support. They also wanted information about the individual consultant or doctor who would be treating them. The government has committed to provide choice of consultant-led team, which suggests the need for information at individual level – information that is not currently available, except in relation to cardiothoracic surgeons.

However, we also found that people did not have stable preferences about what is important to them when choosing a hospital. Preferences – in this case, what influenced hospital choices – were not fixed, but constructed 'on the spot' as participants acquired information (Lichtenstein and Slovic 2006). This suggests there is an opportunity for information providers to influence what people pay attention to by making some aspects of hospital care, such as safety or quality indicators, more salient. Decision aids, or scorecard designs that elicit preferences and values, may in fact help people to identify what is important – that is, to shape their preferences.

Key lessons on types of information

- Presenting too much information means much of it will be ignored.

- Summative measures are helpful for some people, but may be confusing for others.

- The reasons for missing data need to be communicated.

- Information needs to be relevant to the scorecard/procedure.

- Where possible, make the risk information personal to the individual.

- Present data on hospitals not organisations, and, if possible, by consultant-led team.

- Increase the salience of quality and safety information by 'educating' people before they make their choice.

- Recognise that information providers shape preferences, they do not just elicit them.

Presentation of information

The way information is presented can make a difference to how it is used. The way in which information is presented affects people's judgements (eg, inferences about prices or quality), the amount of attention they pay, and also their decisions (Valenzuela and Raghubir 2010). Simple changes will improve understanding, and reduce the burden of cognitive processing. The task of understanding the information needs to be as easy as possible, because the main cognitive task is using this information to form an opinion about each option and then make a choice.

We implemented a number of improvements to the design of the scorecard used in the online experiment, based on our focus groups and previous research:

- We used clear, easily understood labels for indicators that did not require drilling down for further information.

- We made consistent use of intuitive symbols and labels (we used ticks, not traffic lights or a colour-coded scheme).

- We applied evaluative labels (eg, 'poor', 'good', 'excellent').

Labelling values can improve comprehension. Pairing numerical performance information with non-numerical labels was found to increase use of information by less numerate people in experiments around US health plan choices (Peters *et al* 2009). For example, providing evaluative labels such as 'good' and 'excellent' can translate the meaning of numerical information, particularly for those who are less numerate, and ease cognitive processing by conveying meaning about the quality. For example, compare the use of 'poor' to 'below average'. Does being 'below average' necessarily mean that the hospital is performing poorly?

These design features should be adopted by information providers who are presenting comparative quality information for use by the public and patients, particularly for the purpose of choosing between providers. Some of these formats may also be relevant to the presentation of quality information for the purpose of public accountability (for example, in quality accounts), but that was not directly tested in this research.

Key lessons on presentation of information

- Use consistent labelling and symbols across different indicators.

- Consider pairing numerical performance information with non-numerical labels, as this increases use of information by less numerate people.

- Use evaluative labels such as 'good', 'poor' and 'excellent'.

- The indicator labels should use clear language that is easy to understand.

Numerical information

Information about the quality of health care, whether on organisations or individuals, often involves the use of numerical information, and in some cases, such as the Hospital Standardised Mortality Rate, quite complex statistical information.

Our research confirms that high levels of numeracy are required to process this information in order to make a good choice. Given that an estimated 5 million adults in England have a reading age of 11, and 15 million have poor numeracy, many people will

need help to make sense of this sort of information, either through the use of decision aids or an intermediary. It is, perhaps, more likely that these people will rely on the 'information' they currently use to judge quality – that is, their personal experience and that of family and friends, or the recommendation of their GP (Dixon *et al* 2010).

There is a body of evidence which suggests that, even for people with high levels of numeracy, the following guidelines are helpful to improve comprehension.

- When presenting risk information, use 'gain' rather than 'loss' (eg, survival rates rather than mortality rates).

- Values and numbers should be consistent, and if possible, use the same denominator (out of 10 is best).

- Where possible, scaling should be consistent so that higher numbers are always better than lower numbers.

- Use round numbers and avoid decimals.

It is likely that people with different levels of numeracy will require different ways of presenting numerical information. Our research, and previous research carried out by others, suggests that people with higher levels of numeracy:

- prefer specific statistical information

- are more likely to retrieve and use appropriate numerical principles, and to transform numbers presented in one frame to a different frame

- tend to draw more affective meaning from probabilities and numerical comparisons

- want disaggregated information (as do younger people).

People with a lower level of numeracy:

- prefer to receive risk information in verbal or symbolic format

- prefer to receive risk information in frequencies rather than proportions

- like overall summative scores.

Ordering of information

People will spend only a few minutes consulting health information online, so it really matters what information comes first. We designed the online scorecard so that the quality measures were first; the heatmaps confirmed that these were viewed most frequently. We also categorised the indicators into different aspects of quality, such as clinical effectiveness, patient safety and patient experience. Research suggests that people pay more attention when indicators are framed into groups (Fox *et al* 2005). An alternative, if the scorecard is not intended to influence patient preferences by the framing effects, is to have dynamic ordering to prevent over-emphasis of certain indicators.

In our online experiment, we followed the format used on the NHS Choices website, whereby hospitals are presented in columns and indicators along rows. An improved layout, with hospitals along rows and indicators in columns, would encourage option-wise processing, thereby getting people to think through the trade-offs. When you ask people to explain what they are doing when looking at the information, they say they are trying to synthesise it into an overall picture of the hospital. Ordering can help them do this by enabling easy comparison across indicators. Another option (*see* above) is to remove indicators where all hospitals score the same, and therefore make the aspects on which they differ (or differ most) more salient. Furthermore, it is helpful to use a

design which ensures that all the scorecard information is visible on the screen at once – avoiding scrolling, tabs or hyperlinks.

We also thought that by sorting the options (in this case hospitals) by quality, we might also make it more likely that people would choose the highest quality hospital. Although sorting hospitals by quality might appear a good idea, it actually meant that people made worse choices in the first phase of the research. In fact, those who had the options presented by distance made better choices. Clearly, sorting can have perverse effects, and needs to be carefully tested.

There are a number of reasons to explain this counter-intuitive finding. First, it is possible that participants were using their experiences of making other online decisions. On other consumer websites, we are accustomed to higher priced products being in the top rows or on the right-hand side of the screen. In markets for normal consumer goods or services, 'These price inferences translate into quality inferences, and result in items in the vertical and horizontal centres (middle row and/or column) being perceived as a price/quality trade-off and being preferred' (Valenzuela and Raghubir 2010). Second, when items are presented simultaneously, there is a preference for those in a 'central' position rather than extremes (Christenfeld 1995). In our hospital choice experiment, when the options were sorted by distance, the highest quality option was presented in the centre, whereas when sorted by quality, it was the first option viewed. While the central options (third and fourth) were not viewed more frequently (according to the heatmaps), it does appear that there is a tendency to select these options (particularly the fourth). Third, we know that patients consider the local hospital as the 'default' and generally choose this unless there are other factors which mean they would travel further. The local hospital, therefore, is the anchor against which other options are evaluated. By presenting this first, it enabled people to move across the options from left to right, considering how they compared. We cannot say definitively which of these factors was at play, but clearly, sorting of options can have important effects, though these may not always be as predicted.

Key lessons on ordering of information

- Order indicators so that the most important are listed first.

- Group indicators into domains.

- Consider putting indicators in columns, and options in rows, as well as how the information will appear on screen.

- Put the options you want people to select in the middle (at least if the display is horizontal).

- Consider carefully before pre-sorting options.

Use of nudges

The use of nudges has the potential to improve the choices people make, and we had no reason to suspect that they could not deliver benefits if used to help people when faced with a choice of hospital. We were also interested in how nudges affected the experience of choice – that is, how satisfied people were with the choice they made, and whether the nudge had a longer-term impact on people's ability to make informed choices in future.

Our findings suggest that, as with most things in life, no nudge performed well on all these criteria. Of the three nudges we used (pre-select, preference and feedback), the pre-select nudge performed worst in terms of choosing the highest quality hospital. But people who had this scorecard in phase 1 did best in terms of making a choice when presented with a similar task a few weeks later, when nudges were not used. It is possible that people were suspicious of the pre-selected option and intentionally chose something different. They may also have considered other hospitals in relation to the pre-selected option and chosen an alternative – for example, because it was nearer (this is similar to the explanation as to why the quality sort scorecard performed badly). Participants who had the pre-select nudge scorecard were also quite dissatisfied compared with those who had a very simple scorecard (one of the control scorecards, sorted by distance or quality).

A high proportion of those who had the feedback nudge scorecard chose the highest quality hospital. They did slightly worse when presented with a similar task a few weeks later, but were mostly satisfied with the choice they made. Giving people feedback on their selection and asking them to reconsider can increase the proportion that choose a high-quality hospital, but it does not necessarily build their capacity to make better choices in future, without nudges.

Finally, the preference nudge did not do as well as the feedback nudge in terms of the proportion of people making the highest quality choice, but it was better than the pre-select nudge. However, making people think about what was important to them, and making them aware that the quality of hospitals varied, meant that fewer were very satisfied with their choice, and these people did worst when presented with a similar task a few weeks later.

Nudges require the information provider to decide what they want to nudge people towards, and therefore a normative decision is inherent. Here, we chose to nudge people towards the highest quality hospital based on clinical and safety indicators. This might not be politically acceptable, particularly if some of the data on which such judgements are formed are contested. Recent debates about the validity of Dr Foster data on safety suggest that providers are not yet comfortable with judgements being made on the basis of such data, never mind using them to influence patient choices. Perhaps a more acceptable system would be to elicit patient preferences and nudge them towards options consistent with these. However, as our research has shown, patient preferences are inconsistent and unstable, and the very act of presenting information can influence what people think is important when choosing a hospital. Furthermore, the 'best' hospital, on average, may not be the 'best' choice for any particular individual. The ability to personalise information so that decision support can help people choose 'the best for me' is still some way off.

This experiment made use of the ability to present information online and forced people to make their choice, there and then. While online booking is available through Choose and Book, this is not integrated with websites such as NHS Choices. Many providers do not permit direct booking online. However, there is potential in future for there to be a more seamless online service that would allow people to view comparative information and then go directly to book in with their chosen provider.

Our results also show that nudges have different impacts on different people – the feedback nudge made it more likely that older people would choose the highest quality hospital, but in general, younger participants (particularly those with higher levels of education) appear to have been helped more by the nudges. Surprisingly, perhaps, younger participants who had the distance sort scorecard for hospital choice 1 did

particularly badly compared with older participants. Perhaps the latter group's greater familiarity with hospital care and the sorts of issues to consider helped them to choose the highest quality hospital, even though the scorecard was sorted by distance. It appears, from the findings, that younger participants benefited from information about differences in quality of care, and this helped them to understand what is important to them before making a choice.

Key lessons on use of nudges

- Nudges should be used with care, as for some people they may do more harm than good.

- Simply setting the 'preferred' hospital as the default does not make people choose it.

- Making people think about differences between hospitals and what is important to them may increase feelings of awareness and dissatisfaction, choice anxiety, and regret. However, for young people who are not familiar with making these sorts of choices, such information may help them to choose a higher quality hospital.

- Some people may need an advocate to support them in thinking through the options.

- Different nudges suit different people and need to be tested carefully before being widely used to influence patients' choices.

Conclusion

There are a number of methods to improve how comparative information on the quality of hospitals is presented. The coalition government is committed to increasing the availability of information on the performance of hospitals and individuals working within them. Our research reinforces the importance of paying attention to what information is presented and how, and cautions against a mantra that 'more information is always better'.

The 'information revolution' is aimed not only at increasing transparency but is also expected to inform patient choice, and this, in turn, is expected to drive improvements in the quality of care. While other studies have shown that patients are making only limited use of published information, there is an expectation among providers that this will change in future (Dixon *et al* 2010). The same research also found that choice was not yet acting as a powerful lever to drive quality improvement, other than indirectly, through a reputational effect and a desire on the part of providers to retain the loyalty of the local population.

As our research has shown, the information available is complex, and using it to make a decision, even with experimental scorecards, proved challenging. Even though we followed best practice in data presentation to make it easier for patients to interpret the data, all but the highly numerate found the task difficult. Both policy-makers and information providers need to recognise that people find it difficult to make trade-offs between quality, safety, patient experience and location. The government should be cautious about the ability of patients (apart from those who are highly numerate) to make these complex decisions without considerable decision support.

While nudges may help, they are by no means a panacea, and their use needs to be approached with caution given the unexpected and, at times, counter-intuitive effects we observed. More research is needed to evaluate the effects of different nudges on patient decision-making.

Exposing people to the differences in quality between hospitals and forcing them to consider these difficult trade-offs may increase their dissatisfaction with the choice they make. The potential for regret in health care decision-making is great, particularly if you are told that you are more likely to die if you are treated at one hospital than another. Initially, for those unfamiliar with making hospital choices, this might increase the stress and anxiety associated with making a choice.

Policy-makers also need to recognise that selecting a high-quality provider based on clinical quality measures is difficult, and patients may not wish to take responsibility for trade-offs and decisions which they may later regret. Patients may find it easier and less burdensome to make choices on the basis of other factors such as convenience and waiting times, relying on regulators and commissioners to ensure that all hospitals are 'good enough'. Patients may benefit from information which reassures them that hospitals meet required standards. If they feel that any differences they are being asked to consider are differences between 'good' and 'excellent', they may be less concerned than if they feel it is part of their responsibility to make sure that a hospital is not going to kill them. Given these limitations, relying on patient choice as a driver of clinical quality improvement may not be sufficient.

Policy-makers and those responsible for designing web-based information such as NHS Choices should seek to understand the existing research and ensure an evidence-based approach to the public reporting of comparative performance information. Simply allowing all the information currently available on the quality of care to be put into the public domain will not result in people making more informed choices. If the aim of policy is to increase the role of choice in driving quality, the government needs to ensure not only that the data are available, but that those who provide information to patients do so in a way that makes it salient, comprehensible and useful.

Bibliography

Appleby J, Phillips M (2009). 'The NHS: Satisfied now?' in Park A, Curtice J, Thomson K, Phillips M, Clery E (eds), *British Social Attitudes: The 25th report*, pp 25–51. London: Sage.

Audit Commission and Healthcare Commission (2008). *Is the Treatment Working? Progress with the NHS system reform programme*. London: Audit Commission.

Beattie J, Barlas S (2000). 'Predicting perceived differences in tradeoff difficulty' in Weber EU, Baron J, Loomes G (eds), *Conflict and Tradeoffs in Decision Making: Essays in honor of Jane Beattie*. Cambridge: Cambridge University Press.

Bettman JR, Johnson EJ, Luce MF, Payne JW (1993). 'Correlation, conflict, and choice'. *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol 19, no 4, pp 931–51.

Bridgewater B, Keogh B (2008). 'Surgical "league tables"'. *Heart*, vol 94, pp 936–42.

Burge P, Devlin N, Appleby J, Gallo F, Nason E, Ling T (2006). *Understanding Patients' Choices at the Point of Referral*. RAND Technical Report. London: RAND.

Burge P, Devlin N, Appleby J, Rohr C, Grant J (2005). *London Patient Choice Project Evaluation*. London: RAND.

Burgess S, Propper C, Wilson D (2005). 'Will more choice improve outcomes in education and health care? The evidence from economic research'. Bristol: The Centre for Market and Public Organisation (CMPO).

Carlson KA, Bond SD (2006). 'Improving preference assessment: limiting the effect of context through pre-exposure to attribute levels'. *Management Science*, vol 52, no 3, pp 410–21.

Carrigan N, Gardner P, Conner M, Maule J (2004). 'The impact of structuring information in a patient decision aid'. *Psychology & Health*, vol 19, no 4, pp 457–77.

Chaiken S, Trope Y (1999). *Dual-Process Theories in Social Psychology*. New York: The Guilford Press.

Chandon PJ, Hutchinson W, Bradlow ET, Young S (2007). 'Measuring the value of point-of-purchase marketing with commercial eye-tracking-data' in Wedel M, Pieters R (eds), *Visual Marketing: From attention to action*. New Jersey: Lawrence Erlbaum Associates.

Chernev A (2003). 'When more is less and less is more: the role of ideal point availability and assortment in consumer choice'. *Journal of Consumer Research*, vol 30, no 2, pp 170–83.

Christenfeld N (1995). 'Choices from identical options'. *Psychological Science*, vol 6, no 1, pp 50–55.

Coulter A, Ellins J (2006). *Patient-Focused Interventions: A review of the evidence*. London: The Health Foundation. Available at: www.health.org.uk/publications/research_reports/patientfocused.html (accessed on 4 October 2010).

Coulter A, Le Maistre N, Henderson L (2005). *Patients' Experience of Choosing Where to Undergo Surgical Treatment: Evaluation of London Patient Choice Scheme*. Oxford: Picker Institute Europe.

Damman OC, Hendriks M, Rademakers J, Delnoij DM, Groenewegen PP (2009). 'How do healthcare consumers process and evaluate comparative healthcare information? A qualitative study using cognitive interviews'. *BMC Public Health*, vol 9, p 423.

Department of Health (2010a). *Equity and Excellence: Liberating the NHS* [online]. CM 7881. London: Department of Health. Available at: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_117353 (accessed on 4 October 2010).

Department of Health (2010b). *The NHS Constitution for England*. London: Department of Health. Available at: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_113613 (accessed on 22 September 2010).

Department of Health (2009). *Report on the National Patient Choice Survey – March 2009 England* [online]. Available at: www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_103681.pdf (accessed on 24 September 2010).

Department of Health (2008). *High Quality Care For All: NHS next stage review final report*. Cm 7432. London: Department of Health. Available at: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_085825 (accessed on 4 October 2010).

Department for Education and Skills (DfES) (2003). *The Skills for Life Survey. A national needs and impact survey of literacy, numeracy and ICT skills*. DfES Research Report 490. Available at: www.education.gov.uk/research/data/uploadfiles/RR490.pdf (accessed on 24 September 2010).

Diehl K, Zauberman G (2005). 'Searching ordered sets: evaluations from sequences under search'. *Journal of Consumer Research*, vol 31, no 4, pp 824–32.

Dixon A, Robertson R, Appleby J, Burge P, Devlin N, Magee H (2010). *Patient Choice: How patients choose and how providers respond*. London: The King's Fund. Available at: www.kingsfund.org.uk/publications/patient_choice.html (accessed on 24 September 2010).

Dolan P, Hallsworth M, Halpern D, King D, Vlaev I (2010). *MINDSPACE: Influencing behaviour through public policy*. London: Institute for Government.

Dr Foster (2007). 'Hospital guide'. Available at: www.drfosterhealth.co.uk/hospital-guide/ (accessed on 28 October 2010).

Edwards W, Fasolo B (2001). 'Decision technology'. *Annual Review of Psychology*, vol 52, pp 581–606.

Ellins J, McIver S (2009). 'Supporting patients to make informed choices in primary care: what works?' HSMC Policy Paper 4. University of Birmingham: Health Services Management Centre. Available at: www.hsmc.bham.ac.uk/publications/policy-papers/Supporting_patients-PP4-4.pdf (accessed on 5 October 2010).

Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R (2009). 'Public reporting in health care: how do consumers use quality-of-care information? A systematic review'. *Medical Care*, vol 47, no 1, pp 1–8.

Fagerlin A, Ubel PA, Smith DM, Zikmund-Fisher BJ (2007). 'Making numbers matter: present and future research in risk communication'. *American Journal of Health Behaviour*, vol 31, suppl 1, pp S47–S56.

Fagerlin, A, Wang C, Ubel PA (2005). 'Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics?' *Medical Decision Making*, vol 25, no 4, pp 398–405.

Fasolo B, McClelland GH, Lange KA (2005). 'The effect of site design and interattribute correlations on interactive web-based decisions' in Haugtvedt CP, Machleit K, Yalch R (eds), *Online Consumer Psychology: Understanding and influencing behavior in the virtual world*. New Jersey: Lawrence Erlbaum Associates.

Fasolo B, Reutskaja E, Dixon A, Boyce T (2010). 'Helping patients choose: how to improve the design of comparative scorecards of hospital quality'. *Patient Education and Counseling*, vol 78, no 3, pp 273–416.

Fitzsimons GJ, Hutchinson JW, Williams P, Alba JW, Chartrand TL, Huber J, Kardes FR, Menon G, Raghubir P, Russo JE, Shiv B, Tavassoli NT (2002). 'Non-conscious influences on consumer choice'. *Marketing Letters*, vol 13, no 3, pp 269–79.

Foot C, Raleigh V, Lyscom T, Ross S (forthcoming). 'NHS quality accounts: part of the information revolution?' London: The King's Fund.

Fox CR, Bardolet D, Lieb D (2005). 'Partition dependence in decision analysis, resource allocation and consumer choice' in Zwick R, Rapoport A (eds), *Experimental Business Research, Vol III: Marketing, accounting, and cognitive perspectives*. Norwell, MA and Dordrecht, The Netherlands: Kluwer Academic Publishers.

Frank RG, Zeckhauser RJ (2009). 'Health insurance exchanges: making the markets work'. *New England Journal of Medicine*, vol 361, no 12, pp 1135–7.

Gigerenzer G, Todd PM, ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Gourville JT, Soman D (2005). 'Overchoice and assortment type: when and why variety backfires'. *Marketing Science*, vol 24, no 3, pp 382–95.

Grote KD, Newman JRS, Sutaria SS (2007). 'A better hospital experience'. *McKinsey Quarterly*. Available at: www.mckinseyquarterly.com/A_better_hospital_experience_2081 (accessed on 6 October 2010).

Halpern D (2009). *The Hidden Wealth of Nations*. Oxford: Polity Press.

Helsper E (2008). *Digital Inclusion: An analysis of social disadvantage and the information society*. London: Department for Communities and Local Government. Available at: www.communities.gov. uk/documents/communities/pdf/digitalinclusionsummary (accessed on 6 October 2010).

Hibbard JH, Peters EM (2003). 'Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice'. *Annual Review of Public Health*, vol 24, pp 413–33.

Hibbard JH, Peters EM, Dixon A, Tusler M (2007). 'Consumer competencies and the use of comparative quality information: it isn't just about literacy'. *Medical Care Research and Review*, vol 64, no 4, pp 379–94.

Hibbard JH, Peters EM, Slovic P, Finucane ML, Tusler M (2001). 'Making health care quality reports easier to use'. *The Joint Commission Journal on Quality Improvement*, vol 27, no 11, pp 591–604.

Hibbard JH, Slovic P, Jewett JJ (1997). 'Informing consumer decisions in health care: implications from decision-making research'. *The Milbank Quarterly*, vol 75, no 3, pp 395–406.

Hibbard JH, Slovic P, Peters E, Finucane ML (2002). 'Strategies for reporting health plan performance information to consumers: evidence from controlled studies'. *Health Services Research*, vol 37, no 2, pp 291–313.

Hirschman AO (1970). *Exit, Voice, and Loyalty: Responses to decline in firms, organizations, and states*. Cambridge, MA: Harvard University Press.

HM Government (2010). *The Coalition: Our programme for government*. Available at: www.cabinetoffice. gov.uk/media/409088/pfg_coalition.pdf (accessed on 6 October 2010).

House of Commons Public Accounts Committee (2009). *Skills for Life: Progress in improving adult literacy and numeracy – third report of session 2008–09*. Available at: www.publications.parliament.uk/ pa/cm200809/cmselect/cmpubacc/154/154.pdf (accessed on 7 October 2010).

Iyengar SS, Lepper MR (2000). 'When choice is demotivating: can one desire too much of a good thing?' *Journal of Personality and Social Psychology*, vol 79, no 6, pp 995–1006.

Jha AK, Li Z, Orav EJ, Epstein AM (2005). 'Care in US hospitals – the Hospital Quality Alliance Program'. *The New England Journal of Medicine*, vol 353, no 3, pp 265–74.

Johnson EJ, Goldstein DG (2003). 'Do defaults save lives?' *Science*, vol 302, no 5649, pp 1338–9.

The Lancet, editorial (2005). 'Taking health literacy seriously'. *The Lancet*, vol 366, no 9480, p 95.

Lansley A (2010). 'My ambition for patient-centred care'. Speech, 8 June 2010. Available at: www.dh.gov. uk/en/MediaCentre/Speeches/DH_116643 (accessed on 6 October 2010).

Lichtenstein S, Slovic P (2006). *The Construction of Preference*. Cambridge: Cambridge University Press.

Lim JNW, Edlin R (2009). 'Preferences of older patients and choice of treatment location in the UK: a binary choice experiment'. *Health Policy*, vol 91, no 3, pp 252–7.

Lindenauer PK (2009). 'Public reporting and pay-for-performance programs in perioperative medicine: are they meeting their goals?' *Cleveland Clinic Journal of Medicine*, vol 76, suppl 4, pp S3–S8.

Lipkus IM, Samsa G, Rimer BK (2001). 'General performance on a numeracy scale among highly educated samples'. *Medical Decision Making*, vol 21, no 1, pp 37–44.

Lohse GL (1997). 'Consumer eye movement patterns on Yellow Pages advertising'. *Journal of Advertising*, vol 26, no 1, pp 61–73.

Luce MF (1998). 'Choosing to avoid: coping with negatively emotion-laden consumer decisions'. *Journal of Consumer Research*, vol 24, no 4, pp 409–33.

Luce MF, Payne JW, Bettman JR (1999). 'Emotional trade-off difficulty and choice'. *Journal of Marketing Research*, vol 36, pp 143–59.

Marshall MN, Noble J, Davies H, Waterman HA, Walshe K, Sheaff R, Elwyn G (2006). 'Development of an information source for patients and the public about general practice services: an action research study'. *Health Expectations*, vol 9, no 3, pp 265–74.

Marshall MN, Shekelle PG, Leatherman S, Brook RH (2000). 'The public release of performance data: what do we expect to gain? A review of the evidence'. *Journal of the American Medical Association*, vol 283, pp 1866–74.

Mata R, Schooler L, Rieskamp J (2007). 'The aging decision maker: cognitive aging and the adaptive selection of decision strategies'. *Psychology and Aging*, vol 22, no 4, pp 796–810.

Miller GA (1956). 'The magical number seven, plus or minus two: some limits on our capacity for processing information'. *Psychological Review*, vol 63, no 2, pp 81–97.

Moser A, Korstjens I, van der Weijden T, Tange H (2010). 'Themes affecting health-care consumers' choice of a hospital for elective surgery when receiving web-based comparative consumer information'. *Patient Education and Counseling*, vol 78, no 3, pp 365–71.

Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, Zapert K, Turner R (2003). 'The impact of health information on the internet on the physician–patient relationship: patient perceptions'. *Archives of Internal Medicine*, vol 163, pp 1727–34.

NHS Choices website (2010). www.nhs.uk/Pages/homepage.aspx (accessed on 6 October 2010).

NHS Choices website (2009). *Performance Report December 2009*, www.nhs.uk/aboutnhschoices/ professionals/developments/documents/2010/091204%20december%202009%20kpi%20% 20dashboard%20v0%20do1.pdf (accessed on 24 September 2010).

Nicholas D, Huntington P, Williams P (2004). *Digital Consumer Health Information and Advisory Services in the UK: A user evaluation and sourcebook*. Centre for Information Behaviour and the Evaluation of Research. London: City University. Available at: www.ucl.ac.uk/ciber/Summary%20 Report%20260404.pdf (accessed on 6 October 2010).

Payne JW, Bettman JR, Johnson EJ (1993). *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.

Peters E, Dieckmann NF, Västfjäll D, Mertz CK, Slovic P, Hibbard JH (2009). 'Bringing meaning to numbers: the impact of evaluative categories on decisions'. *Journal of Experimental Psychology: Applied*, vol 15, no 3, pp 213–27.

Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK (2007a). 'Less is more in presenting quality information to consumers'. *Medical Care Research and Review*, vol 64, no 2, pp 169–90.

Peters E, Hibbard J, Slovic P, Dieckmann NF (2007b). 'Numeracy skill and the communication, comprehension, and use of risk-benefit information'. *Health Affairs*, vol 26, no 3, pp 741–8.

Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S (2006). 'Numeracy and decision making'. *Psychological Science*, vol 17, no 5, pp 407–13.

PricewaterhouseCoopers (2009). *Champion for Digital Inclusion: The economic case for digital inclusion*. Available at: www.parliamentandinternet.org.uk/uploads/Final_report.pdf (accessed on 7 October 2010).

Raleigh V, Foot C (2010). *Getting the Measure of Quality: Opportunities and challenges*. London: The King's Fund. Available at: www.kingsfund.org.uk/publications/quality_measures.html (accessed on 7 October 2010).

Reutskaja E, Camerer C, Nagel R, Rangel A (forthcoming). 'Search dynamics in consumer choice under time pressure: an eye-tracking study'. *American Economic Review*.

Robertson R, Dixon A (2009). *Choice at the Point of Referral: Early results of a patient survey*. London: The King's Fund. Available at: www.kingsfund.org.uk/publications/choice_at_the_point.html (accessed on 7 October 2010).

Rosen DL, Olshavsky RW (1987a). 'A protocol analysis of brand choice strategies involving recommendations'. *The Journal of Consumer Research*, vol 14, no 3, pp 440–44.

Rosen DL, Olshavsky RW (1987b). 'The dual role of informational social influence: implications for marketing management'. *Journal of Business Research*, vol 15, no 2, pp 123–44.

Rosen DL, Olshavsky RW (1985). 'Use of product-testing organizations' recommendations as a strategy for choice simplification'. *Journal of Consumer Affairs*, vol 19, issue 1, pp 118–39.

Ryan RM, Deci EL (2000). 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being'. *American Psychologist*, vol 55, no 1, pp 68–78.

Salvucci DD, Goldberg JH (2000). 'Identifying fixations and saccades in eye-tracking protocols'. *Proceedings of the 2000 Symposium on Eye-Tracking Research and Applications*. New York: ACM (Association for Computing Machinery), pp 71–78.

Schkade DA, Kleinmuntz DN (1994). 'Information displays and choice processes: differential effects of organization, form, and sequence'. *Organizational Behavior and Human Decision Processes*, vol 57, no 3, pp 319–37.

Schlesinger M (2010). 'Choice cuts: parsing policy-makers' pursuit of patient empowerment from an individual perspective'. *Health Economics, Policy and Law*, vol 5, no 3, pp 365–87.

Schneider EC, Lieberman T (2001). 'Publicly disclosed information about the quality of health care: response of the US public'. *Quality in Health Care*, vol 10, no 2, pp 96–103.

Schwartz B (2004). *The Paradox of Choice: Why more is less*. New York: Harper Perennial.

Schwartz B (2000). 'Self-determination: the tyranny of freedom'. *American Psychologist*, vol 55, no 1, pp 79–88.

Shanteau J, Thomas RP (2000). 'Fast and frugal heuristics: what about unfriendly environments?' *Behavioral and Brain Sciences*, vol 23, no 5, pp 762–63.

Shekelle PG, Lim Y, Mattke S, Damberg C (2008). *Does Public Release of Performance Results Improve Quality of Care? A systematic review*. London: The Health Foundation. Available at: www.health.org. uk/publications/research_reports/performance_results.html (accessed on 6 October 2010).

Spranca M (2000). 'Designing websites to empower health care consumers' in Weber EU, Baron J, Loomes G (eds), *Conflict and Tradeoffs in Decision Making: Essays in honor of Jane Beattie*. Cambridge: Cambridge University Press.

Stanovich KE, West RF (2000). 'Individual differences in reasoning: implications for the rationality debate?' *Behavioral and Brain Sciences*, vol 23, no 5, pp 645–65.

Stevenson FA, Kerr C, Murray E, Nazareth I (2007). 'Information from the internet and the doctor–patient relationship: the patient perspective – a qualitative study'. *BMC Family Practice*, vol 8, p 47.

Thaler RH, Sunstein CR (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven and London: Yale University Press.

Thorlby R, Maybin J (2010). *A High-Performing NHS? A review of progress 1997–2010*. London: The King's Fund. Available at: www.kingsfund.org.uk/publications/a_highperforming_nh.html (accessed on 6 October 2010).

Tversky A, Kahneman D (1974). 'Judgment under uncertainty: heuristics and biases'. *Science*, vol 185, no 4157, pp 1124–31.

Ubel PA, Jepson C, Baron J (2001). 'The inclusion of patient testimonials in decision aids'. *Medical Decision Making*, vol 21, no 1, pp 60–68.

Uhrig JD, Harris-Kojetin L, Bann C, Kuo TM (2006). 'Do content and format affect older consumers' use of comparative information in a Medicare health plan choice? Results from a controlled experiment'. *Medical Care Research and Review*, vol 63, no 6, pp 701–18.

Vaiana ME, McGlynn EA (2002). 'What cognitive science tells us about the design of reports for consumers'. *Medical Care Research and Review*, vol 59, no 1, pp 3–35.

Valenzuela A, Raghubir P (2010). 'Are top-bottom inferences conscious and left-right inferences automatic? Implications for shelf space positions'. Working paper. www.econ.upf.edu/docs/seminars/ valenzuela.pdf (accessed on 7 October 2010).

Walton J, Berkowitz E (1979). 'The effects of choice complexity and decision freedom on consumer choice behavior' in Wilkie W, Abor A (eds), *Advances in Consumer Research: Vol 06*. Association for Consumer Research.

Wensing M, Grol R (2005). 'Educational interventions' in Grol R, Wensing M, Eccles M (eds), *Improving Patient Care: The implementation of change in clinical practice*. Edinburgh: Elsevier.

Wilson TD, Lisle DJ, Schooler JW, Hodges SD, Klaaren KJ, LaFleur SJ (1993). 'Introspecting about reasons can reduce post-choice satisfaction'. *Personality and Social Psychology Bulletin*, vol 19, no 3, pp 331–39.

Wilson TD, Schooler JW (1991). 'Thinking too much: introspection can reduce the quality of preferences and decisions'. *Journal of Personality and Social Psychology*, vol 60, no 2, pp 181–92.

Zuckerman M, Porac J, Lathin D, Smith R, Deci EL (1978). 'On the importance of self-determination for intrinsically motivated behavior'. *Personality and Social Psychology Bulletin*, vol 4, no 3, pp 443–46.