

Exploratory Interaction and Extended Cognition: Redesigning Decision-making Support in Healthcare

Dissertation with a Practical Component in Media

Presented to Goldsmiths, University of London

In Fulfilment of the Requirements for the Degree of Doctor of Philosophy

Supervisors

Prof. Matthew Fuller – Centre for Cultural Studies

Prof. William Gaver – Department of Design

Martin Feuz

Centre for Cultural Studies / Department of Design

Goldsmiths, University of London

Declaration

I declare that the work presented here is, to my best knowledge, original unless otherwise acknowledged in the text.

Martin Feuz

Zürich, October 2014

Abstract

This thesis presents practice-based research on redesigning decision support within the area of prostate cancer screening. More fundamentally, this research is specifically interested in a politics and practice of supporting exploratory interactivity in healthcare decision-making by non-experts. The motivation for this research stems partly from insights gained from my own empirical research into algorithmic web-search results personalisation as well as issues underlying the biopolitical logic of cancer screening practices, the individualisation of risk and the profound uncertainties inherent in evidence-based medicine which are largely underarticulated in the design of decision-support tools for individuals. Taking these conditions as the driver for this research, the first and theoretical part of the thesis analyses models of interactivity and the politics of statistically-derived medical knowledge in evidence-based medicine as well as problematising dominant but narrow conceptions of human decision-making. Following such analysis are the insights that the statistical nature of epidemiological risk information is of limited applicability for individual decision-making and, thus, patient preferences matter to guide decision-making under uncertainty. By introducing cognition understood as socially distributed and extended into and performed through the environment, this research proposes to rethink how to design for exploratory information interaction in medical decision support. Following a research-through-design method and developing a minimal reserved design approach, a number of prototypes were developed to investigate the potentials hypothesised in the theoretical part of the thesis. The prototypes assumed a probing function in shedding further light on the thinking and practices of medical professionals, which leads me to suggest repositioning them as *probotypes*, occupying a middle ground between prototypes and cultural probes. Ultimately, the contribution of this research lies in critically and practically exploring conceptual and methodological potentials for redesigning exploratory interactions in shared medical decision-making processes.

Acknowledgements

This work was made possible through the support of numerous people. First, I would like to extend my warmest gratitude to my supervisors Matthew Fuller and Bill Gaver for their inspiration and continuous intellectual guidance. It has been a long, productive and challenging journey through unexpected waters. Without their support, this journey would have been unimaginable.

A special thank you to Felix Stalder who has set me on this journey many years ago over a few cups of coffee. I also want to thank Nathalie Jeremijenko and Marsha Rosengarten for stimulating conversations and inspiration.

A thank you goes to Cédric Poyet from the Prostate Carcinom Center at the University Hospital Zürich, Felix Huber of the Medix Group Practice as well as Jacques Cornuz and Reto Auer from the University Hospital Lausanne for their interest in this work and valuable conversations.

Further thanks for intellectual conversation and friendly support go to Aki, Antonio, Chuck and Bill, Cornelia, Daniel, Dimi, Janina, Kione, Lukas, Marc, Niccolo, Raphael and Simon.

And a very special thanks to Renato and Stefano, my extended family including Alex, Yorkie, Jacky, Darius for making me feel at home while commuting and in particular to Monya who inspired and supported me and, many years ago, kicked this all off. Lastly, a warm thank you to my family for being patient while I was working on this.

The thesis is dedicated to my brother Markus.

Table of Contents

List of Figures

List of Abbreviations

Introduction

Chapter 1 **16**

The problem with health web-search interaction

Chapter 2 **30**

Medical epistemology - the history of statistics and its role in medical evidence

Chapter 3 **77**

Shared medical decision-making and decision support

Chapter 4 **94**

Prospective reflection on issues in prostate cancer screening decision support

Chapter 5 **121**

Focus of practiced-based research and operationalization

Chapter 6 **131**

Design project and method

Chapter 7 **172**

Evaluation and reflection on research findings

Chapter 8 **193**

Discussion of implications and future directions

Bibliography

Appendix 1: NHS PSA decision aid

Appendix 2: Prototype 1 - Venn Interaction Diagram

Appendix 3: Prototype 2 - Reflection Cards

List of Figures

Chapter 1

Figure 1.1: Standard model of search process

Chapter 2

Figure 2.1: Diagrammatic overview of position of evidence-based medicine

Figure 2.2: Relation of Health and Illness in pre-Broussais phase

Figure 2.3: Relation of Health and Illness in Broussais phase

Figure 2.4: Relation of Health and Illness in risk factor phase

Figure 2.5: Canguilhem's perspective on health and normalcy

Figure 2.6: Tukey's view on process of research idea and question generation

Chapter 3

Figure 3.1: Visualising Uncertainty about the future

Chapter 4

Figure 4.1: Overdiagnosis rates of cancers

Chapter 6

Figure 6.1: Prostate cancer screening decision process overview

Figure 6.2: Grouped visualization style of prostate cancer screening outcome

Figure 6.3: Idea #1b - Step 1: select position

Figure 6.4: Idea #1b - Step 2: positive PSA test result

Figure 6.5: Idea #1b - Step 3: negative biopsy result

Figure 6.6: Historical view of relations of health and illness

Figure 6.7: Idea #2a - mapping of PSA test result

Figure 6.8: Idea #2a - mapping of PSA test result - epidemiological perspective

Figure 6.9: Idea #2a - mapping of PSA test result - individual perspective

Figure 6.10: Idea #2a - mapping of PSA test result - movable scale

Figure 6.11: Conceptual design for interactive Venn diagram

Figure 6.12: Interactive Venn diagram - minimum healthy and pathological zones

Figure 6.13: Interactive Venn diagram final prototype - epidemiological perspective

Figure 6.14: Interactive Venn diagram final prototype - individual perspective

Figure 6.15: Prospective reflection cards - first prototype

Figure 6.16: Prostate cancer screening process and informational gaps

Figure 6.17: Projection of reflection cards for decision-making process

Figure 6.18: Distribution of cards under assumption of positive PSA test result

Figure 6.19: Early card prototype and updated one

Figure 6.20: Different sizes of card prototypes

Figure 6.21: Reflection card final prototype, back and front

List of Abbreviations

BMI	Body Mass Index
CDC	Centers for Disease Control (USA)
DA	Decision Aid
EBM	Evidence-based Medicine
GFT	Google Flue Trends
HCI	Human Computer Interaction
ILI	Influenza-related Illness
PSA	Prostate Specific Antigen
RCT	Randomised Controlled Trial
SDM	Shared Decision Making
SEO	Search Engine Optimisation

Introduction

This thesis is interested in the politics and practice of health information interaction. This focus has partly originated from my earlier empirical research on online search interaction and the ways in which statistical methods are employed in filtering and providing individually relevant search results (Feuz et al. 2011). While in some circumstances, such as for example when trying to re-find a previously visited website, such filtering may indeed be very helpful, using online search engines for health search queries, a knowledge domain often unfamiliar to those undertaking the search, seems very problematic. This is because the ways in which such filtering occurs is obscured and not available for further inspection or sense-making. In turn, the person cannot assess the relevance of the search results she is being presented with and the kinds of assumptions based on which relevancy is supposed. I speculate that such an approach is not supportive of making sense of health information within the context of health search interactions, in particular since it has come to light that for the majority of people health information found online affects their decision-making (Zickuhr 2010, 3). Furthermore, the model underlying the design of such interactions seems unfit to support a more exploratory type of interaction and provide the relevant means to do so. Given these problems, I have chosen to approach them with an encompassing review of literatures identified as relevant for medical decision support with the goal to uncover connected, specific and significant aspects of such processes to be conceptually rethought and redesigned.

From exploring and analysing the history and more recent issues with medical epistemology as well as contemporary approaches towards decision support, I have chosen to focus on developing tools for the navigation of probabilistic risk information around prostate cancer screening as a topic for this thesis and practice-based research project. This focus is relevant as an interesting and productive case study for this research project for the following reasons:

Firstly, for the past four decades, evidence-based medicine (EBM) has slowly but increasingly become dominant as a knowledge paradigm and clinical practice approach in western medicine. EBM is the “use of mathematical estimates of the risk of benefit and harm, derived from high-quality research on population samples, to inform clinical decision-making in the diagnosis, investigation or management of individual patients” (Greenhalg 2010, 1). While, on the one hand, EBM has certainly contributed to making the medical evidence base more robust, on the other hand, it has also fallen victim to issues such as selective reporting and reporting bias which distort the

evidence base significantly (Goldacre 2012; Ioannidis 2005).

Secondly, and independently of the issues identified concerning the reliability of EBM, the applicability of epidemiologically-derived cancer risk factors for deciding on cancer screening at the level of an individual person is questionable. This is because framing epidemiological risk factor information and thus probabilities derived from a sample population as relevant for individual decision-making, rather than public health management only, obscures the significant levels of uncertainty inherent in such information. This is despite the inability to know any one individual's health outcome (Esposito 2007; Politi et al. 2007, 683). Unfortunately, the probabilistic nature of epidemiologically-derived risk information is not well understood by decision making patients (Gigerenzer 2002, 4; Gigerenzer and Gray 2011, 3). Furthermore, the uncertainties underlying such medical risk information and subsequent decision-making are seldom made transparent by medical professionals (Politi et al. 2007, 690).

Thirdly, shared decision-making has been suggested as a relevant means to engage with preference-sensitive care situations. These are healthcare situations, such as prostate cancer screening, in which several treatment options are considered to have equivalent health outcomes when measured in terms of health economics indices, such as life expectancy and quality of life months. Treatments, however, affect patients differently due to their side-effects. Thus, as Mulley et al. argue (2012), patient preferences matter in such situations. Unfortunately, medical professionals do not appear to be very good at diagnosing patient preferences (*ibid.*), which brings me to the last point.

Decision aids often come in the form of videos or pamphlets and are developed to support people in making an informed and unbiased decision in a medical situation. This is thought possible by providing a balanced view of benefits and harms of treatment options according to the current evidence base. The use of decision aids as part of shared decision-making processes has been successful in some ways (Cochrane Collaboration 2009). As meta-analysis by the Cochrane Collaboration indicates, when compared to common care interventions, decision aids performed favourably by improving people's knowledge about options available and continue to dampen "enthusiasm for major elective surgery [...] in favour of more conservative options," as well as accounting for a decreased preference for prostate specific antigen (PSA) testing (*ibid.*, 43). A PSA test is a blood test widely applied in prostate cancer screening to identify potential risk of cancer. The test is highly controversial (Ablin 2010). The review results, however, also point out that decision aids do not perform better than conventional care practices (that is consultations

without the support of decision aids) when it comes to satisfaction with decision-making and anxiety.

I argue that the above key finding may be due, among other aspects, to the fact that the interactional model underlying the design of decision aids assumes that patient preferences and values are pre-established and only need to be retrieved by decision-makers when confronted with a decision to be made about screening. Instead, Epstein and Peters (2009, 195) argue that in medical situations characterised by substantial uncertainties and potential “outcomes that have previously not been considered or cannot be imagined,” preferences are constructed rather than simply elicited. Unfortunately, the preference construction process is not very well supported as part of shared decision-making processes today. Furthermore, the uncertainty underlying the probabilistic nature of cancer risk information in decision aids is exclusively presented as natural frequencies. While natural frequencies may improve readability and avoid certain biases, I argue that such an approach simultaneously further obscures the fundamental uncertainties underlying the generation and applicability of such information, in particular at the level of the decision-making individual.

The critical review and analysis of relevant literatures in chapters 1 to 4 has been conducted in order to elicit a multiperspectival framing of the problem. This includes online health information search and digital culture studies, social and cultural studies of evidence-based medicine, theories of cognition and design research. In the practice-based research, I will work on the problem of decision-support from these perspectives and how they provide different facets to the problem. To this end, the insights gathered from the critical analysis operate as a set of connected problems and thus thinking resources which my practice-based research attends to. Thus, this research draws inspiration, both from the phenomena of actual health information practice but equally from theoretical reflections such a health information practice is constituted and affected by.

Following from these considerations, the question of how to design for an understanding of unpredictability and uncertainty within the context of prostate cancer screening decision-making is an interesting one. The central hypothesis for this research is that shared medical decision-making within the selected context could be productively supported by an exploratory model of interaction. I will argue that such a model of interaction is better suited to circumstances of sense-making and decision-making under uncertainty and, thus, to the aspirations of a patient-centred care seeking to deeply respect patients as unique living beings and care for them on their own terms (Epstein and Street 2011). The contribution of this research lies in critically and

practically exploring conceptual and methodological potentials for redesigning exploratory interactions in shared medical decision-making processes. More specifically, this research project has identified and problematised the role of risk and uncertainty in cancer screening decision-making, its potential grave implications for patients as well as starting to redesign and evaluate prototypes with medical professionals in this regard. The particular contributions are as follows.

Firstly, through the critical analysis of relevant literature, a set of problematic aspects in contemporary health information interaction was identified. These contribute relevant facets from which the problem of redesigning for an understanding of risk and uncertainty in cancer screening decision-making could be engaged. In turn, this provides helpful orientation on how to reconceptualise decision-support.

Secondly, these facets helped defining a set of design requirements informing the design research. These include a reconsideration of what constitutes a useful and necessary scope and time frame relevant for facilitating the decision-making process in its entirety and complexity. Furthermore, the desirability to convey a differentiated understanding between an epidemiological and individual perspective of risk as regards uncertainty involved in such risk information and decision-making. This also includes, the particular ways in which risk and uncertainty are articulated (natural frequencies) and visually represented (magnitudes of risk). Lastly, the need for supporting personal preference construction and, thus, reconsider patient autonomy beyond the idealised decision-making individual was identified.

Thirdly, two prototypes were developed to address these design requirements. The first one is an interactive Venn diagram helping to understand false-positive and false-negative test results by providing a means for exploring the meaning of quantitative PSA test results. The second prototype consists of a set of reflection cards, which aim to facilitate the personal preference construction process.

The prototypes were evaluated with medical professionals with regards to their topical relevance, clarity, productivity and meaningfulness in addressing decisional problems as identified in the critical analysis. The medical professionals found the prototypes useful in many ways. They thought them well designed and declared an interest to evaluate them, including actual patients. They also highlighted potential limitations in that the reflection cards may not appeal to some patients, not are such cards able to unbiased doctors with a set opinion regarding the merit of prostate cancer-screening.

Fourthly, unexpectedly and importantly, my prototypes came to function as

probes for diagnosing the underlying and practiced understanding of and dynamics between medical professionals and their roles. This leads me to suggest repositioning them as probotypes, occupying a middle ground between prototypes and cultural probes.

Overview of the thesis

Chapter one starts with key observations on how online search engines have come to be used for finding health information online and the scale and scope at which this occurs. In light of the massive growth of online content over the last two decades, a brief historical development of search engines will provide some explanation for their dominant operative mechanisms such as the algorithmic filtering of supposedly relevant results as well as the interaction paradigm that has developed. I argue that algorithmic filtering is particularly problematic for exploratory search interactions where the search goal may not always be very clear from the beginning. I introduce and discuss exploratory search and interaction as opposed to an information-retrieval model of interaction. In the remainder of the chapter I consider other conceptions of interface and interactivity, which will serve as further inspiration and set the orientation for this research on exploratory health interaction more broadly.

Chapter two is concerned with better understanding the body of knowledge for which exploratory interaction aims to be productive. Specifically, I will analyse and discuss western medical epistemology and the development of evidence-based medicine, including issues such as reporting bias and selective reporting that have recently surfaced. Epidemiology and thus statistics are among the key underlying modes in which such knowledge is generated, hence my discussion in more detail of their historical development as well as their contemporary application via cancer screening. Whilst epidemiologically-derived risk information may be said to be beneficial for public health management, they simultaneously introduce a significant dilemma for individual decision-making. Furthermore, I will consider biopolitical concerns around evidence-based medicine (EBM), such as the changing possibilities and responsibilities projected onto people in sustaining and promoting their health. The chapter will end with a reflection on notions of normalcy and pathology, which seem particularly strained in the context of the probabilistic means by which epidemiological thresholds for normalcy and pathology are determined and applied for purposes of cancer screening. This specific topic will set the focus for considering approaches to decision support in chapter three.

In **chapter three**, I build upon the concerns discussed in the previous

chapter by considering the assumptions underlying the design of decision aids as a tool for promoting shared decision-making (SDM) for cancer screening. Decision aids (DA), in this context, aim to help people in making an informed and unbiased decision by providing a balanced view of benefits and harms of treatment options according to the best current medical evidence base. The chapter will define SDM, consider its genealogy and discuss the principles based on which decision aids are designed and evaluated. This renders more transparent the underlying assumptions and design decisions materialising in the design of contemporary DA's. In particular, I will analyse the specific ways in which risk information is presented and the implications this has on individuals' sense- and decision-making. Risk visualisations typically privilege the display of magnitudes of risk at the cost of the unpredictability of risk factors. I argue that, from the perspective of an individual person, this seems partly a questionable design decision. This is because communicating and understanding unpredictability is at least as important as it suggests preparing for a variety of possible health outcomes. This relates to the second part of the chapter dealing with the role patient preferences and values play in shared decision-making for cancer screening. Some suggest that medical practitioners should focus on helping patients cope with uncertainty rather than just understanding it (Politi et al. 2007, 690). One way in which this could be achieved is by supporting the preference construction process.

Chapter four and five will consolidate the critical review of the literature in chapters one to three and prospectively reflect on the issues identified from the perspective of and potential for redesigning decision support. I will argue that the contemporary design approach to decision support is largely based on an information-retrieval model of interaction which tends to obscure fundamental issues concerning personal preferences underlying the decisional problem in preventive cancer screening. Instead, I hypothesise that an exploratory model of interaction is better suited to circumstances of sense-making and decision-making under uncertainty and thus to the aspirations for patient-centred care (Epstein and Street 2011). In the second part of the chapter I will introduce the extended mind thesis and concepts of distributed cognition, guided by which I will investigate the productive potential of an exploratory model of interaction to support shared decision-making. In **chapter five**, I discuss the selection of issues from the critical analysis which will be addressed and operationalized in the practice-based research.

In **chapter six**, I document and discuss the iterative design process. Research through design (RTD) (Gaver 2012, 940), the method used in this research, is a method which privileges the exploration of a research question

through the articulation and evaluation of designs. I discuss the successive development of two final prototypes. The first prototype is focused on supporting an understanding of false-positive and false-negative prostate specific antigen (PSA) test results. This derives from the critical analysis in chapters 1 to 4 regarding the limitations of using epidemiological risk information in decision-support at the individual level as well as from the finding that risk information visualisations often represent magnitudes of risk rather than the actual random distribution of risk at the individual level (see chapter 3). In aiming to develop epistemic interactivity for an understanding of false-positive and false-negative test results, this design research aims to render more intelligible and meaningful both the specific meaning of probabilistically derived risk information as well as its limited predictive ability at the individual level.

The second prototype is concerned with supporting the personal preference construction processes in preference sensitive medical care situations, such as in cancer screening. As discussed in the critical analysis, this is relevant because decision-support based on epidemiological risk factor information alone does not help prepare for coping with the uncertainties inherent in the screening process and the ways in which its implications affect individuals in their personal and social lives. As a consequence, preferences and values matter when making decisions about cancer screening. The design process is inspired by Mol and Law's (2004) suggestion to privilege action over knowledge. Without negating knowledge, their perspective is interested in the implications of medicine for the lived practice of being a patient and what we might learn from this for medical practice. The prototype consists of a set of reflection cards containing scenarios a person might encounter when conducting prostate cancer screening. The scenarios are crafted in such a way that they embed the experience and potential implications of screening in everydaylife settings, be that in relation to family life, relationships, work environments or leisure activities. The reflection cards aim to support reflection on how people engaged in cancer screening would react in the described scenarios and who they would turn to for advice and help. As identified in the critical analysis, this reconceptualises patient autonomy beyond the individual decision-making person and builds on distributed processes of meaning-making.

To this end, the design research approach followed is carried out by offering a set of prototypes that address problems identified in the literature review. In **chapter seven** I discuss the evaluation process with medical professionals, and in **chapter eight** I evaluate the research findings. Together, the design research contributes to the critical analysis in the following ways. Firstly, it has

identified aspects that matter when designing for an understanding of uncertainty of PSA test results as opposed to probabilistically-determined risk, which is the focus of contemporary decision aids. The design research has also been productive in discovering ways in which shared decision making could be furthered by more actively involving the participating patient through epistemic interaction in the sense-making process, and how the design may best support this. Furthermore, the design research has brought to light that the actual screening path contains much more friction due to the need of undertaking multiple biopsies so as to avoid false-positive and false-negative PSA test results. Some patients find it hard to accept negative biopsy results and thus the absence of cancer when a PSA test result has previously indicated the contrary. This is an important aspect that the literature review has not brought to light. It matters significantly in light of this research, which aims to foreground potential psychological and sociocultural implications the cancer screening experience may have for people. Lastly and most importantly, the design research contributes to the critical analysis in that it has unearthed an understanding of the professional roles and expectations among the medical professionals. This is relevant as in some cases it might affect the role decision-support interventions such as those I have designed might come to play in actual clinical decision-making practice.

Finally, a number of productive lines of inquiry are shown indicating strong potential for developing this work in useful and relevant ways.

Chapter 1

The problem with health web-search interaction

Introduction and research context

According to the Pew Research Center, searching for health information online had become the third most common activity among internet users aged 18 and above in the USA by 2010 (Zickuhr 2010, 3). For many, health information found online affects their medical decision-making in terms of how to treat an illness or ask doctors new questions (ibid., 4). However, there are significant issues with understanding and carefully assessing such health information. Among the issues identified are, for example, difficulties in making sense of statistical information (Gigerenzer 2002, 4; Gigerenzer and Gray 2011, 3), lacking a capacity to critically assess the quality of medical information (Kukla 2007, 33) and a variety of suggested biases which can have negative implications for sense- and decision-making (see chapter three). Furthermore, in the context of cancer screening, significant levels of overdiagnosis and overtreatment have been identified, that is the “diagnosis of a ‘cancer’ that would otherwise not go on to cause symptoms or death” and which may nevertheless be treated preventively (Welch and Black 2010). Prostate and breast cancer screening programmes and corresponding diagnostic tests have recently been critically reappraised and, as a consequence, are no longer recommended in some countries.¹

Within the context of prostate cancer, which is the focus of this research project, given the unknown effects of early detection with a PSA test in terms of promoting longevity and reducing morbidity as well as the risk of serious side-effects (BMJ Group), the decision of whether or not to screen is a difficult one. It requires understanding relevant medical risk information and preventive treatment options, possible implications following such courses of action as well as considering personal values and preferences. Unfortunately, the probabilistic nature of epidemiologically-derived risk information is not well understood by decision makers (Gigerenzer 2002, 4; Gigerenzer and Gray 2011, 3). Furthermore, the uncertainties underlying such medical risk information and subsequent decision-making are seldom made transparent by medical professionals (Politi et al. 2007, 690).

Due to the novelty of the situation, values and preferences have typically

¹ Screening for Prostate Cancer, UK National Screening Committee, 2010. Stellenwert des PSA-Wertes bei der Früherkennung des Prostatakarzinoms, Swiss Medical Board, 2011.

not been previously deliberated and, thus, may not be available to guide medical decisions (Epstein and Peters 2009, 195). Instead, as Epstein and Peters (ibid.) suggest, they need to be constructed first. This involves understanding the potential benefits and harms of treatment options and, more importantly, the ways in which this affects a person's life beyond bodily functions. The preference-construction and deliberation process may be supported by physicians or influenced by family and friends. Such preference-construction processes, however, are not well supported by shared medical decision-making processes (Epstein and Street 2011, 102), which seem chiefly focused on the provision of best medical evidence in a balanced manner by way of decision aids. The aim of this research is to critically reflect on the limitations and implications of evidence-based medicine in relation to decision support for prostate cancer screening decisions. Such reflection will inform the redesign of decision support in the practice-based part of this research, which will also include the preference-construction process.

In summary then, making an informed decision on prostate cancer screening is characterised by multifarious complexities bearing on the information environment, the doctor-patient relationship, and potential decisional conflict in the absence of guidance by personal values and preferences. These complexities are what renders this an interesting and relevant case study. This introductory chapter begins to explore some of these complexities by considering the ways in which health search interactions with contemporary web search engines are supported. This is relevant to consider because search engines affect health decision making (Zickuhr 2010, 4), operate as mediators as discussed in this chapter, and, in that regard, act similarly to decision aids within the context of shared decision-making for prostate cancer screening, as will be discussed below and in further detail in chapter three. I argue that both of them take an active and mediating role as part of the interaction with them. As Fuller (2012, 134) argues with regard to decision-support systems more broadly

the interest of the decision support system lies less in the quality of the advice it provides than in its redistributive function, in the way that it can help displace and redefine expertise, valid knowledge, the landscape of choice and the rationalisation one makes of it.

A brief analysis of the historic development of web search engines allows for an understanding of the dominant interaction model of search interactions, the operative mechanisms underlying and supporting such a model, as well as the implications this might have for information needs and decision tasks within the context of health information search. As I analyse and argue in chapter three, the interactional model underlying both search engines and decision

aids in shared decision-making is largely based on an information-retrieval understanding of interaction and, thus, less suitable within the context of cancer screening. In contrast, in this research project I argue for and design towards exploratory ways of interacting with health information. Such exploratory interactions are hypothesised as relevant for supporting an understanding of the limitations of the probabilistic nature of epidemiologically-derived health risk information as well as supporting people in exploring and constructing preferences and values for guiding their decision-making. Before analysing the issues concerning the relevance of epidemiological risk information for individual decision-making in chapter two as well as the relevancy of personal preferences within the context of cancer screening in chapter three, a short history of search engines will be discussed. This will help understand how the dominant interaction paradigm with search engines has developed, and, in turn, allow to draw parallels with the contemporary design of decision aids in shared health decision-making as discussed in chapter three.

A short history of search engines

Search engines have played a vital role in the development of the World Wide Web (web). This is mainly due to the fact that the web lacks an inherent indexing and categorising mechanism. From early on Yahoo and other commercial operators offered a directory to the web, which was compiled by human experts in a fashion similar to a library catalogue. While this format was helpful in accessing information on the web, it simply could not cope with the rapid growth of the web. Subsequent search engine providers started to build automated indexing and ranking mechanisms. Indexing is the part of a search engine where search crawlers continuously update the search engine's index by scanning the web for new websites and for new content on those indexed websites. Ranking, on the other hand, is concerned with the matching of a user's search query with the index and, based on a set of rules, presents the user with a set of search results. At the core of ranking lies the trade-off between precision and recall. Precision, on the one hand, is concerned with the accuracy of the match between search query and retrieved search results. Recall, on the other hand, is concerned with number of relevant search results produced. When increasing precision of the match between search query and results, recall is reduced and vice versa.

While better suited to cope with the enormous growth of the web, early automated search engines suffered from a number of problems. Key among

these was that their ranking mechanism chiefly relied on a relatively crude statistical keyword-matching process between search query and indexed webpages. Thus, with the growth of content on the web, this produced an enormous amount of search results. More often than not, it would require a strenuous effort to find useful search results. This was particularly true for exploratory search behaviour, where the search goal is not always very clear and emerges as part of the search process. This is a type of search which will be at the core of my focus and which will be considered in more detail shortly.

A further problem search engines have faced from early on up to this day is bias in search results by advertisers. With the introduction of e-business and the advent of a potentially global marketplace, it became particularly opportune to be listed as a relevant vendor in the top rankings of search engine results. Thus, a new industry was born: Search Engine Optimisation (SEO). Service providers in this industry basically help businesses to improve their listing position. This ranges from ensuring that a business Website is properly indexed with the relevant keywords to manipulating their ranking position by applying various methods inflating their relevancy. The former practices work along the methods suggested and approved by search engines for organic listing, while the latter aim to influence the search rankings beyond the methods approved by search engines.

Characteristics of search engine results such as comprehensiveness, scalability, relevancy, objectivity and spam emerged as the core challenges search engine operators had to deal with, and remain so today. Larry Page and Sergey Brin aimed to address these problems when they introduced the Google search engine (1998). As they explained in their paper (1998) the system architecture of Google was from the outset planned to scale along the growth of the Web. Their paper suggested that this could be achieved by focusing on fast crawling technology, efficient use of indices storage, a capacity of the index system to compute large data sizes as well as scaling at the user end by handling search queries quickly.

The key differentiator, though, was their proposal for “bringing order to the Web” (ibid.) with the introduction of the PageRank algorithm based on link analysis. The link analysis concept was born out of structural similarities between the hyperlinked structure of the Web and bibliometrics, that is the study of the importance of a piece of academic writing for its field on the basis of citation structures. Given the Web’s hyper-linking structure and its resemblance to academic writing, the concept seemed equally well-suited. Link Analysis or PageRank, as Google’s founding fathers Page and Brin came to name it, applies this concept to the Web. PageRank is a method based on algorithms that measure the importance of a Webpage by considering the

number of in-links to this Webpage from other Webpages and, more importantly, the importance of the inlinking Webpages. The importance of these Webpages was again determined by their own PageRank. In other words, importance is determined by means of popularity within a chain of linked Webpages. This approach appears to be democratic in nature, as Google points out in their description of PageRank below:

PageRank relies on the uniquely democratic nature of the Web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for the importance of page B. But Google looks at considerably more than the sheer volume of votes or links a page receives; for example, it also analyses the page that casts the vote. Votes cast by pages that are themselves 'important' weigh more heavily and help to make other pages 'important.' Using these and other factors, Google provides its views on pages' relative importance.

(Google Technology Page 2005)²

Interesting, however, is the fact that some pages are in “themselves important” and weigh more (ibid.). While it may be argued that some hierarchy seems in place, there appear to be no transparent means available to understand what is of a priori importance and what is not. At best, an a priori unimportant Website can gain importance over time through popularity. Compared to existing search engines at the time, Google soon managed to gain significant market share because their search engine produced much quicker results which seemed more relevant, up-to-date and less advertising-biased.

The interface has not changed very much since the early days: the almost empty page with a search box and the mostly ten blue links has become a quasi-standard. However, while the front-end does not seem to have changed much, the back-end has changed in a number of ways, in some areas substantially so. Search results for a given search query are nowadays automatically filtered by a number of variables that remain hidden in the immediate user interaction. The goal of this is to further increase the relevancy of search results. Among others, search results are typically filtered based on the search user's geo-location as derived from her IP (Internet Protocol) address. Thus, if a search user uses Google and enters “restaurant” as a search query, the search engine assumes that the user is looking for a restaurant within the city she currently accesses the Internet from. This may obviously be useful in some cases, but largely depends on context.

² <https://web.archive.org/web/20050315015924/http://www.google.com/technology/> (accessed March 20, 2014).

However, a more substantial and difficult to identify change to the ranking mechanism is that some time ago Google started to personalise users' search results, promising to deliver more relevant results to the user whose query is now being considered in the context of her search history and other data previously compiled into a personal profile (data shadow). In order to produce this context, vast amounts of personal information need to be collected, organised and made actionable. Within the fast receding limitations of storage space and computing power, profiles can never be too comprehensive, too detailed, or too up-to-date. Google is compiling personal profiles in three dimensions: the knowledge person (what an individual is interested in, based on search and click-stream histories), the social person (whom an individual is connected to, via e-mail, social networks and other communication tools) and the embodied person (where an individual is located in physical space, and the states of the body) (Stalder and Mayer 2009). Together, these three profiles promise to provide a detailed, comprehensive and up-to-date context for each search query, with the potential to deliver precise results that reflect not just the information 'out-there' but also the unique interest a user has at any given moment.

Personalised search does not simply aim at providing a view onto existing reality (Introna and Nissenbaum 2000). Rather, it promises an 'augmented reality' in which machine intelligence interprets the user's individual relationship to reality and then selects what is 'good.' As a result, it has become unlikely that two users see the same search results for a particular search query even when accessing from the same IP address (Feuz et al. 2011). Search engines, thus, often act as black-boxed mediators, both on the level of mere functionality as well as on the data set, without articulating such mediation as part of the interface. Unfortunately, the use-context in these cases is determined denotatively and without consent from the user. Such practices are deeply problematic as these interactions are pre-empted and operate on an unknown operational model of the world (Holmes 2007; Stalder and Meyer 2009). Unfortunately, many search engine users do not seem to be aware of this (Pan et al. 2007). This seems hardly surprising given that they are not made aware of these mechanisms operating in the background as the search interface and search results do not indicate any of these operations. But to fully understand its implications it is necessary to adopt a more nuanced focus and reflect on it in light of different types of search interaction behaviour and today's typical search engine interface.

Broder (2002) suggests differentiation between three different types of search behaviour, which is a useful taxonomy for the purposes of the present discussion. Firstly, navigational search queries, where users want to find the

URL for a specific website. Secondly, transactional search queries, such as checking flight prices, which can be performed on a number of different yet specific websites. Thirdly, informational search queries to find information that may be present on multiple websites and where the search goal may not always be clear at the beginning but only emerges through the search process itself. From this perspective it becomes clearer that search results personalisation is especially problematic regarding informational search queries. Thus, the analysis and argument I develop will mainly focus on this search behaviour.

Exploratory search interaction

Recently, the notion of informational search behaviour has been developed further and characterised as 'exploratory search' (White et al. 2007), which is the name I will use from now on. Exploratory search interactions are characterised by a number of typical features. To start with, very often there is a complex informational problem at hand and a desire to learn about it. Also, people engaging in exploratory search may be unfamiliar with the knowledge domain their search goal relates to. This may include a lack of understanding of dominant and peripheral actors within that domain as well as of useful search approaches. Furthermore, people may not have good knowledge of keywords, concepts and information sources which may be relevant to formulate search queries and evaluate search results. Lastly, it is possible that exploratory searchers may not have a specific search goal in mind when they set out. The goal may only evolve through a process of learning about the specific knowledge domain, its concepts and actors within it. Given these characteristics, the exploratory search process typically develops over the course of multiple sessions, which may last days, weeks or months.

White and Roth (2009, 6) mentions people "grappling with chronic illness, work teams creating complex solutions or products, learners studying complex material over time, families making long-term plans, scientists investigating complex phenomena, and hobbyists tracking developments over a life-time" as people with an exploratory search interest. It is hardly difficult to imagine and extend this list to include conspiracy theorists, journalists, engaged citizens, young parents determined to be knowledgeable and responsible about their child's development, activists, teachers, relatives of people with a health condition, immigrants to a country organising their lives, environmentalists, politicians and many more. In short, there may be a multitude of reasons why search engine users may well be interested in going

beyond the top ten 'most relevant' search results, be it for personal development, curiosity and learning, deciding on intricate social or environmental issues or making better-informed decisions within the context of prostate cancer screening, which is the focus of this research project.

Having explicated the characteristics of exploratory search interaction, I will now reflect on today's search engine interfaces and the practice of personalisation in light of exploratory search interactions. This is relevant for better understanding the mediating role of search engines within the context of exploratory health search interaction. Instead of supporting exploration and critical appraisal of the information found online, the hierarchical top-listing of search results authors their supposed high relevancy (Pan et al. 2007; Rogers 2009). Many search engine interfaces typically build on a commonly accepted set of action grammars and handles suggested by the Human Computer Interaction (HCI) domain. The action grammar applied more often than not aims at describing a context-free metasyntax and, thereby, suggesting universal applicability and usability. In the case of universal search engines, such as Google or Bing, as mentioned previously, the interface is typically made of a single search box with a search button on an otherwise almost empty page. A user enters her search query, clicks on the search button and is then presented with the 'ten most relevant search results' for that search query. Typically, the user then has a few very general refinement options available to further narrow the search. What remains hidden are the aforementioned numerous assumptions at work that lead to the ranking and filtering of those ten most relevant search results.

This approach arguably works well for simple navigational tasks in web search. However, with exploratory search tasks, this dominant search interaction model poses a number of problems. As highlighted in the introduction of exploratory search behaviour above, the process involves a number of steps which require a much more symbiotic form of interaction with the search engine. The process involves information seeking, filtering, reading and learning as well as sense-making. In exploratory search, the search goal is typically unclear at the beginning, thus highly precise search results for a potentially not yet well formulated search query may be less useful than a broader coverage of a topic. As an exploratory searcher sets out on a search task, becoming familiar with a novel knowledge domain is important and involves learning about concepts and actors in that space. Considering today's search results pages, hardly any support is available to familiarise oneself with a domain. Typically, the search results page indicating the millions of relevant search results found presents the top ten of them and allows inspecting related search keywords. While there are a number of other filters

offered, such as document type filters, time-range and reading level, it seems questionable how these should provide much help in coming to better understand a knowledge domain. Furthermore, today's search interaction does not support well search processes that may last multiple sessions and involve collaboration with others. This is particularly problematic with regard to the personalisation of search results. The possibility that everyone gets different search results for the same search query makes it more difficult to explain and relate to others how one has found a specific piece of information or why one has not found it.

Interaction models

Today's search engine interfaces can be described as relying on a few core assumptions that resonate with an information-processing model of the mind. Broder's (2002) standard model of the search process illustrates this well.

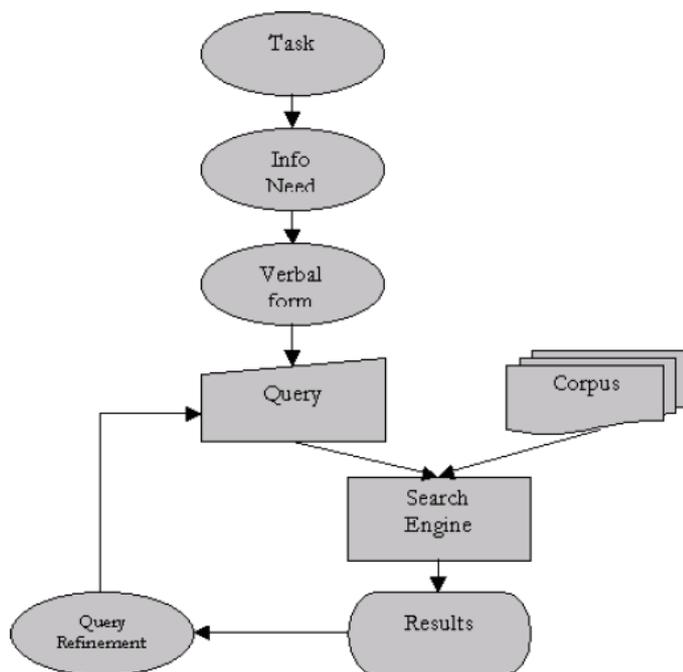


Figure 1.1: Standard model of search process

The model presents a highly abstracted and de-contextualised view of interaction. Cognitive processes here may be said to be represented as purely mental processes consisting of “identifying an information need, followed by the activities of query specification, examination of retrieval results, and, if

needed, reformulation of the query, repeating the cycle until a satisfactory result set is found” (Hearst 2009). Alternate models of the search process such as Bates’ berrypicking have made very useful contributions to a more interactive style of search by suggesting more iterative processes of searching, learning and the shifting of focus and goals.

While the model itself seems rather formal, Bates (2002) recently clarified her perspective by suggesting a need to include biological and anthropological aspects for a fuller understanding of information seeking. In his keynote at the 2008 European Conference on Information Retrieval, Nicholas Belkin, a distinguished Professor and long time researcher on Information Retrieval, suggested with great concern to move “beyond the limited, inherently non-iterative models of IR that we have been concerned with, to the development of models of IR which incorporate the user as an active participant in the IR system” (Belkin 2008). Among the challenges ahead in developing the field further, he argues that “understanding and supporting information behaviours other than specified search” is of great importance but little so far is known about it. Furthermore, he continues with criticising mainstream IR research for its prime concern with efficiency and effectiveness and its subsequent neglect of the role of affect, in particular since Kuhlthau has demonstrated as early as 1991 its relevance. He concludes his keynote by proposing that this may require “to give up the idea of strictly formal models of IR in favour of realistic and useful models of IR. This, in my opinion, may not be a bad trade off” (Belkin 2008).

Unfortunately, even in the more recently developed research area of exploratory search, most work is based on such understandings, as the review of the evaluation methods considered by White suggests (2009). However, given the specific characteristics of the exploratory search process described earlier, I argue that a different model of interaction as well as cognition may better support and inform the design of such interactions. Thus, in summary, both the typical highly black-boxed character of search engine interactions as well as the more recent shift towards personalisation of search results are generally problematic, particularly concerning exploratory search queries, where the capacity to gauge the relevancy and sources of search results may be more limited if not absent. As we will see in chapter three with regards to the role and productivity of health decision aids, I argue that the absence of supporting exploratory ways of interacting with health information is detrimental to shared health decision-making and, thus, patient autonomy.

Given the narrow commitments to an understanding of what makes up interaction and the relevant environment, it seems useful to consider another body of work reflecting on the ways of interacting. This will serve as inspiration

for this research on exploratory types of interaction as well as informing and orienting the practice-based study as conceptualised and set-out in chapters five and six.

Interface and Interactivity reconsidered

The role of software interfaces has also been analysed within the realm of software studies. Fuller (2003) discussed the explanatory power of metaphors applied in software interfaces and transferred from other knowledge domains. In so doing he identified that, while initially suitable to overcome bootstrapping problems, they tend to lose their explanatory power as software is functionally upgraded and, thus, outgrows such literal comparison. This merely symbolic representation of computational events further obfuscates computational processes and maintains a clear demarcation as to possibilities for users to interact with the system in unanticipated but nevertheless useful ways. Fuller demonstrates this through the analysis of Sennett's example of a (re-)visited Chicago bakery, which has been equipped with a 'user-friendly' interfaced baking system. The system allows the mostly temporary workers to bake all sorts of bread by the simple touch of a button. Unfortunately, "much of the product is wasted when the actual temperature, the rising of the loaves, or some other factor fails to match the representation on screen. There is a skip outside full of burnt loaves, victims of automated friendly fire" (Fuller 2003, 106). One has to assume that, according to the designers of the baking systems, "workers controlling the process via the interface have no need for an understanding of how to bake bread. The process is illegible to them" (ibid.). Similarly and earlier, Suchman (1987) has articulated the shortcomings of those narrowly-conceived models of working processes and the ways in which these are mapped onto interfaces, privileging certain work processes while ignoring the possibility to adapt these depending on situational need.

Dourish (2004) has also argued for reconsidering the model of interaction based on which much of HCI work has developed. He follows a phenomenological approach based on Husserl, Heidegger and Merleau-Ponty as well as considering the work of Suchman on situated action. I will discuss aspects of his arguments relevant to the exploratory search interaction process. One important aspect of Heidegger's phenomenology was the notion of Dasein, which he has introduced to overcome Husserl's earlier mentalistic understanding of intentionality.³ For him Dasein, or being-in-the-world, was the

³ Broadly speaking, 'intentionality' refers to "directedness" of meaning, that is how a meaningful

essence of how we encounter the world and make sense of it; intentionality, thus, is inseparable from being and doing.

Dourish illustrates his concept with the example of the hammer. We sometimes act through technology (ready-to-hand) and sometimes are conscious of the tool (present-to-hand) when we need to adjust a tool. For him the tool comes into being through the transition of ready-to-hand to present-to-hand, thus appreciating a process of meaning-making as practical and purposeful. For Dourish (2004, 138), “intentionality sets up a relationship between embodied action and meaning.” What we make of a tool in terms of meaningfulness seems intricately linked to action and context. In order to make intentional references effective, Dourish stresses the role of coupling, by which he means how the intentional reference can be maintained and managed. For example, coupling is at work when a blind person uses a cane through which she establishes and manages a sense of orientation. While the intentionality is on sensing the environment, this is made effective through the blind cane which operates as an extended sensory prosthesis and, according to Heidegger, withdraws from being present-at-hand to being ready-to-hand when coupling is effective.

In the case of web search, the ‘search’ button as the tool is not much available for adjustments to coupling. Typically, a search user evaluates the search results and, if need be, reformulates the search terms in an iterative process of search as highlighted in the formal models of web search information retrieval above. In fact, the identified user behaviour of ‘thrashing’⁴ seems to be indicating some ‘breakdowns’ in tool-use, particularly since the tool, i.e. the search engine and its interface, does not render itself transparent in how it works and, thus, for the coupling to be available for adjustments.

Dourish (2004) builds on the work of Suchman (1987) to detail how meaning is not only inseparable from action but also often situated in a task and social context. Suchman demonstrated how technology is embedded in communities of practice that operate with and through distinct, ad-hoc and situated organisations according to task context. Rather than developing interactive technology according to pre-formulated and generic plans, she suggested to consider what earlier ethnomethodologists Garfinkel and Sacks (1970) have named accountability, or the observable and reportable character

relationship between two entities is set up. Earlier philosophers such as Brentano and Husserl conceived intentionality as the purely mental phenomena, and, thus, different from the physical action in the world (Dourish 2004, 136).

⁴ Thrashing occurs when users continuously search with mostly similar keywords, thus exhibiting an anchoring bias. As a consequence, typically search results will not change much and, thus, leave users unsatisfied (Morville and Callender 2010).

of practiced rationality (Suchman 1987). For Dourish (2004) this allowed her to demonstrate how “problems with interactive technology lay in the imbalance between the situated organization of practical action and the regimented models that systems embody.”

Much of the work on situated action and accountability has focused on work environments such as the computer-supported cooperative work area. Little work exists on search interactions analysing their exploratory and socially situated nature. However, this seems of vital importance as web search engines clearly have become a vital social infrastructure. Thus, how one makes sense of and accounts for the web search results found seems important to allow for multiple subjectivities. Recent research on the use of and reliance on highly-ranked search results (Pan et al. 2007) has highlighted that search users tend to trust artificially higher-ranked search results more than those ranked lower on the search results page, despite the fact that their abstracts seemed less relevant to the task. Thus, accountability for web search results, in the absence of specific domain knowledge, seems mostly to rely on the status authoring role through those search engines (Rogers 2009). Considering such experimental evidence in light of personalisation algorithms and exploratory health information search processes indicates an urgent need to research and design alternate and differently engaging means of searching and making sense of the results provided. As will be discussed further in the conclusion as well as developed in more detail in chapter two, being able to find the most ‘relevant’ health information online (this chapter) or robust medical evidence (chapter two) is certainly important. Nevertheless, the ways in which this is or is not applicable, useful or meaningful for individual decision-making (chapter three) constitutes a significant and relevant element of exploratory health information interaction. As I argue in chapter two and three, current approaches fail to account for this. The ways in which this can be better supported will be explored theoretically and practically in this research project.

Conclusion

To summarise, while online health information interaction is very popular and, according to some research (Fox 2006), leads to mostly positive experiences, I argue that the model underlying web search interactions is not always suitable to the task of online health information search. This is because the dominant model presumes that a person knows exactly what she is searching for as well as having well-developed skills to assess relevant health information found. As

research indicates, the search result position itself modulates trust and relevancy, despite seeming semantically less relevant (Pan et al. 2007). The exploratory search process is further obscured by algorithmic filtering mechanisms, targeting search result relevancy based on a user's prior search and click history (data shadow). Within the context of a novel medical diagnosis, such data shadows may seem to be of little relevance if not potentially misleading.

Within the context of prostate cancer screening decision-making, as I analyse and argue in chapter three, the model underlying interaction with search engines and decision aids in shared decision-making is largely based on an information-retrieval understanding of interaction. In contrast, due to the nature of some health information search and decision tasks described above and discussed further in chapter three, I suggest that an exploratory interaction model would be more suitable and supportive. More specifically, exploratory interactions are hypothesised as productive in supporting an understanding of the limitations of the probabilistic nature of epidemiologically-derived health risk information for decision-making by individuals in cancer screening (chapter two) as well as supporting people in exploring and constructing preferences and values for guiding their decision-making (chapter three).

Chapter 2

Medical epistemology - the history of statistics and its role in medical evidence

Introduction

The previous chapter has looked at interaction with search engines and models of search interactivity in particular. By doing so it aimed at shedding light on the assumptions underlying the specific conceptualisations of search interactivity and their implications. This chapter is devoted to the medical knowledge domain and, more specifically, the area of medical knowledge production and distribution. The recent proliferation of novel data-driven approaches towards medicine is believed to have vital implications for what constitutes medical data and the methods by which such data is then used in generating medical knowledge (O'Reilly Media 2012). This chapter, thus, seeks to a) generate a deeper understanding of the different kind of knowledge that novel data approaches such as Big Data might be said to produce, and b) assess the implications of medical decisions based on them.

In order to develop such an assessment, this chapter will first examine contemporary mechanisms of production of knowledge in evidence-based medicine (EBM) and distribution via surveillance-medicine (Armstrong 1995). As will become evident, much of contemporary western EBM is based on probabilistic statistical methods. Unfortunately, as it becomes increasingly apparent, there are deep systematic flaws in the existing medical evidence base as a consequence of various incentive structures in academia and commerce that engender selective reporting and publication bias (Goldacre 2012; Ioannidis 2005). Furthermore, whilst probabilistic methods may be said to be beneficial for public health management – which, as the genealogical section will demonstrate, is also where they originated – they simultaneously introduce a significant dilemma for individual sense- and decision-making regarding normal and pathological states of health.

Thus, before being able to understand and assess the kinds of ways in which novel data-driven approaches conceptualise and define normalcy and pathology, my genealogical analysis will unravel the biopolitical (Foucault 1973) aspects of western evidence-based methods of production and distribution of medical knowledge. This analysis will be further substantiated by examining the relevant historical development of key statistical methods used in medical knowledge production, such as randomisation and statistical significance tests. In addition, this chapter will demonstrate that the current

data explosion is not the first one in the history of medicine and, in light of developments in the life sciences and bioengineering, it seems unlikely to be the last one. I will consider, map and analyse relevant historical developments to compare them to current data paradigms and evaluate them. Thereby, I will discuss the application of the probabilistic approach to ever finer degrees of risk factor-based diagnosis of anticipated, near and actual pathological conditions as well as the problematic implications this has for self-surveilling, decision-making individuals. As the genealogical section identifies, epidemiological risk factor and quantitatively-determined pre-pathological diagnosis are based on the assumption of normalcy and pathology neatly sitting on a linear gradient, which renders them measurable and convincing at least across populations. In turn, such an assumption provides the necessary condition to allow for anticipatory regimes of pre-pathology to be probabilistically argued in the first place. I will challenge this assumption on the basis of an updated version of Canguilhem's posited concept of normalcy and pathology that involves a much longer-term consideration informed by evolutionary theory (Canguilhem 1994).

In light of the identified implications of the biopolitics of contemporary medical knowledge production, I will analyse the mechanisms of Google Flu Trends that represent a Big Data-driven approach to medical knowledge production. I will argue that, from an epistemological perspective, Big Data should be seen as an intensification of existing probabilistic methodologies, such as those of epidemiology discussed below, rather than a categorical shift. To this end, it will further the focus on statistically significant patterns and, as a consequence, render individual discrepancies and the individual less important and desirable to attend to. Historically, the applicability of averages in medicine was seen as immoral, as successful treatment of some individuals based on such averages comes at the cost of death for those unaccounted for therein (Porter 1988, 157). At the level of the decision-making individual, whether as doctor or patient, such statistical modulations, rather than mere informational backdrop for rational examination, become performative as active mediators. My aim is to cast light on the ways in which such performativity is facilitated through the presentation and distribution of probabilistic risk information.

Lastly, as described and argued in the preceding chapter, the current paradigms of search interaction are defined by a dominant understanding of human cognition following an information-processing model of mind. This chapter starts establishing a link between the ways in which search interactions are conceptualised, the presumed autonomy and self-explanatory capacity of data discussed in this chapter and the design of decision-making

structures and systems discussed in the next chapter.

In the following chapter, I will specifically consider the underlying cognitive assumptions regarding the decision-support technologies in the area of shared medical decision-making processes. Underlying conceptions of search interactivity, evidence-based medical knowledge production and decision-support technologies are based, I argue, on a specific and constrained conception and understanding of human cognition. In contrast, as I will set out in chapter 4, it is promising to make such a link explicit and consider alternate processes of cognition as a productive force. This is hypothesised to be particularly fruitful in the context of redesigning exploratory healthcare information interactions. The view I will propose is one of socially distributed and environmentally extended processes of cognition and mind. Such a perspective, I hypothesise, generates a productive environment based on which to sketch and evaluate inventive and proliferating means for supporting healthcare decision-making processes.

Evidence-based medicine

As will be discussed throughout this chapter, western medicine today is increasingly practiced via the evidence-based medicine (EBM) approach. The following section will define, outline and examine its characteristic knowledge production and distribution mechanisms. It will highlight the way in which data and statistical methods have become its fundamental underpinning principles and illustrate how they operate. Furthermore, I will identify on the basis of recent examples various problematic implications of EBM and discuss how they affect practitioners and patients alike. I will argue that the issues identified can be mapped against a set of underlying assumptions in statistical thinking and epidemiology whose genealogical becoming I will attempt to trace. As I will emphasise, early developments in statistical thinking and political attitudes of the medical profession at the time had impacted significantly on medical epistemologies. In significant parts this took place through the forces discussed below that have led to the development of EBM and the ways in which EBM as an approach to the health of people has also influenced contemporary advances in the life sciences, e.g. in genetics and neuroscience. In this sense, medical history will be examined here not as a trajectory of gradual, linear development. Rather, and in the spirit of Foucault (1976), an attempt will be made to

Diagnose the conditions based on which the linkages between medicine and contemporary reality were formed [and] trace out the diverse connections and liaisons that brought it into existence and gave it its salience and characteristics.

(Rose, in Jones and Porter 2001, 50)

Following such an approach will be useful in uncovering the assemblage of assumptions underlying the various fields, such as medical epistemology, statistics, medical ethics, decision research and cognitive science, that interweave in shaping shared decision-making processes.

For the past four decades, evidence-based medicine (EBM) has slowly but increasingly become dominant as a knowledge paradigm and clinical practice approach in western medicine. This is evident in the forms of new institutions, such as the Cochrane Collaboration and in the UK the National Institute for Clinical Excellence (NICE),⁵ new journals, recurring editorials in leading medical journals as well as the adoption of EBM-methods, such as randomised controlled trials in mainstream medical research (Timmermans and Berg 2003). Archie Cochrane was an epidemiologist pioneering the field in the 1970's and founder of the Cochrane Collaboration that defines EBM as

Evidence-based health care is the conscientious use of current best evidence in making decisions about the care of individual patients or the delivery of health services. Current best evidence is up-to-date information from relevant, valid research about the effects of different forms of health care, the potential for harm from exposure to particular agents, the accuracy of diagnostic tests, and the predictive power of prognostic factors.

(Cochrane Collaboration 2014)

Trisha Greenhalgh (2010,1), a Professor of primary health care and researcher interested in sociology and medicine, more recently added to the above the “use of mathematical estimates of the risk of benefit and harm, derived from high-quality research on population samples, to inform clinical decision-making in the diagnosis, investigation or management of individual patients.” Evidence-based clinical practice is a sub-discipline that constitutes how EBM is practiced with patients. Muir Gray (1997, 20), another early pioneer in the

⁵ The National Institute for Clinical Excellence is a UK-government funded body with the goal to reduce variation in the availability and quality of NHS treatments and care (http://www.nice.org.uk/aboutnice/howweare/who_we_are.jsp; ac. 07.03.2013).

EMB field, defines this as “an approach to decision-making in which the clinician uses the best evidence available, in consultation with the patient, to decide upon the option which suits that patient best.”

These definitions are useful as a starting point as they set out the goals and central means of production and distribution of the evidence-based medicine approach as well as the attitudes towards patients. The next sections will take up and discuss the following important notions in the above definitions: ‘predictive power of prognostic factors;’ ‘mathematical estimates of the risk of benefit and harm;’ ‘population samples’ and ‘best evidence in consultation with patient to decide which option suits patient best.’

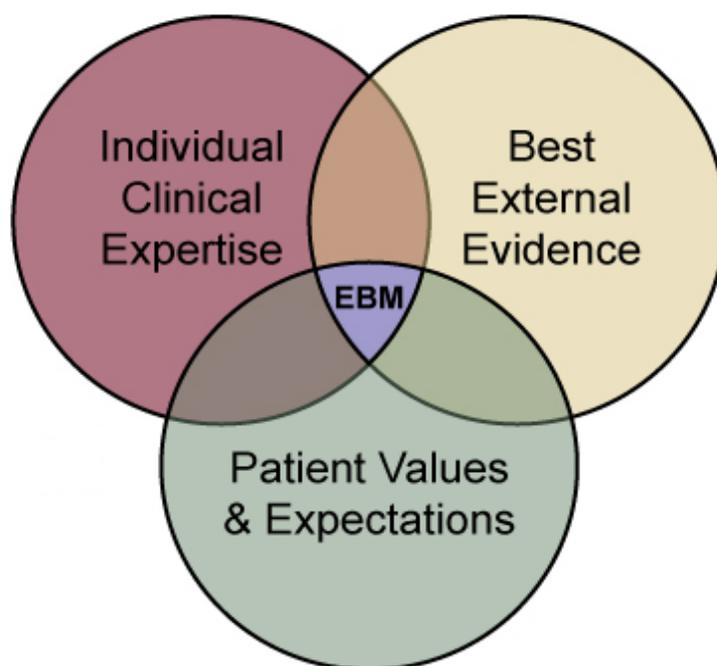


Figure 2.1: Diagrammatic overview of position of evidence-based medicine

Firstly, as highlighted in the definitions above and illustrated in Figure 3, EBM builds on existing evidence which is generated by clinical research in observational studies and clinical trials. The findings from known studies and trials on a topic are then aggregated and systematically reviewed. The goal of a systematic review is to collate and consolidate relevant and high-quality research on a specific research question and minimise bias by employing statistical methodological rigour, such as controlling for various potential biases (for example selection bias⁶ and study effect sizes⁷), to ensure internal

⁶ Selection bias is a consequence of non-random sampling errors, which may lead to certain members of a population to be over-emphasised (Jadad and Enkin 2007).

statistical validity and clinical significance of findings. Such evidence is then used by clinicians to assess the risks and benefits of various treatments.

A short history of evidence-based medicine

In order to better understand how and why evidence-based medicine exists as it does today it is useful to consider its brief history. Until about 40 years ago, physicians in clinical practice, it was thought, would make medical decisions regarding their patients based on ‘clinical judgment.’ This meant they would “synthesise all of the important information about the patient, relevant research, and experiences with previous patients to determine the best course of action” (Eddy 2005, 10). However, in the early 1970’s Wennberg, a medical care epidemiologist, identified stark geographical discrepancies in clinical practice for what was considered essentially identical medical problems (Chassin et al. 1986; Wennberg 1973, 1982).

In order to better understand and explain such discrepancies, these observations were examined more closely from a working hypothesis of clinical judgment under uncertainty. The assumption here was that physicians cognitively struggle to assemble and assess all the relevant information when consulting a patient (Eddy 1982, 249). When investigating how physicians diagnosed diseases in patients as well as making sense of evidence based on laboratory equipment, such as x-rays under controlled conditions, wide variations were identified, including cases when doctors re-examined the same material on two separate occasions.⁸ Interestingly, some have argued that the variation in clinical practice was partly due to uncertainty arising in the minds of doctors as a consequence of the expansion of medical technology and the difficulty in interpreting test results of such machinery (Eddy 1984, 76). Issues of judgment and decision-making are important for this research and will be looked at in more detail in the following chapter. The implications of such significant variation in clinical practice were to be found in quality of treatment which, in some cases, may have been inappropriate or even unnecessary, hence costly in terms of risk, pain, distress and money (Chassin et al. 1986; Dartmouth Atlas 2007; Wennberg 1982).

While part of the early interpretations for the disconcerting variations in

⁷ Effect size is measuring the strength of the relationship between two variables in a statistic (Jadad and Enkin 2007).

⁸ There is a growing body of literature inter- and intra-observer ‘variation’ which can be easily identified when conducting relevant literature research online.

clinical practice focused on uncertainty in the judgment and decision-making process of medical practitioners, such uncertainty did not only arise from limitations in practitioners' capacity to cognitively synthesise relevant information and exercise proper judgment. In addition, and more importantly for an understanding of medical evidence, a second line of argument identified a lack of systematisation and rigorous verification in existing processes of generating and applying medical knowledge as potentially implicated in variations in clinical practice. Such shortcomings could explain why uncertainty and variation in the individual decision-making process on the part of practitioners occurred in the first place.

In his highly cited short book *Effectiveness and Efficiency: Random Reflections on Health Services*, Cochrane (1972) identified a number of issues characterising the state of medical evidence and practice. Firstly, medical evidence from clinical research was not systematically tested, collated and consolidated.⁹ By that Cochrane referred, for example, to the introduction and routine clinical adoption of cervical smear tests for preventing carcinoma of the cervix before rigorous testing via RCT. Conducting a RCT after widespread clinical adoption was considered unethical, as it would have necessitated that a control group would only receive a placebo. Secondly, pre-established standard therapies and tests created ethical dilemmas, as applying a randomised controlled trial (RCT) to test their efficacy would involve denying treatment to half of the patients (*ibid.*, 23). Lastly, with regards to diagnosis, he observed a lack of standardisation of procedures in diagnostic laboratory tests as well as disinterest by clinicians in such precise results (*ibid.*, 38). Others have added that clinical guidelines were based on individual authors' opinion rather than collated research (Eddy 1990; 2005, 11).

As becomes rapidly obvious from the above discussion, at the core of both EBM and the observed variations in clinical practice lies the more detailed question of what are considered normal and pathological states of a patient as well as what is considered good evidence to diagnose such states. Thus, EBM must also be seen as an attempt to standardise such definitions. The following section will discuss the ways in which EBM approaches the issue of evidence. As medical evidence becomes operative in defining and delineating normal from pathological states, this chapter more broadly then provides genealogical insights in order to understand the various developments of concepts of evidence with regards to normalcy and

⁹ For Cochrane (1972, 27), this led, for example, to the introduction of cervical smears testing procedures hoping to prevent cervix carcinoma into clinical practice that was previously not rigorously RCT-tested.

pathology. This will set the stage for evaluating more novel data-driven approaches in such a context at the end of this chapter.

Mechanisms of western medical knowledge production

To understand the body of knowledge EBM aims to assess by systematic reviews, it is necessary to examine the key mechanisms of knowledge production in clinical medical research. This includes assessing the safety and effectiveness of medications, treatments as well as medical equipment. Importantly, clinical research studies follow a population studies approach and, thus, rely on statistical techniques, such as randomisation, significance and hypothesis testing. Randomisation as a scientific method was proposed by 19th century Belgian astronomer Quetelet (1842) and will be explained further below in relation to randomised controlled trials and discussed in more detail in the genealogical section on data and statistics. Significance and hypothesis testing methods were developed in psychology in the work of mathematician and pragmatist philosopher C.S. Peirce and further elaborated and promoted by R. A. Fisher as well as J. Neyman and E. Pearson.

Fisher was a British statistician and evolutionary biologist with a keen interest in genetics as well as eugenics. His early work, though, was concerned with field experiments in agriculture and led to the development of the statistical techniques mentioned which he laid out in his book *Statistical Methods for Research Workers* (1925). A significance test is a statistical method for assessing the likelihood of experimental results being due to chance occurrence. Fisher proposed it as an informal index to be used as part of the non-quantifiable process of drawing conclusions from observations (Goodman 1999a). It calculates the probability of obtaining the experiment results under the assumption that the item tested, for example a new treatment, has no effect. Results within the arbitrary but now widely standard 5% cut-off rate are typically considered as statistically significant and open the doors for academic publication (Dwan et al. 2008; Sterling 1995). In fact, reviews found that more than 95% of articles in psychology journals claim statistical significance (Sterling 1959; 1995). Such publication behaviour is deemed statistically improbable¹⁰ and, as a systematic review confirms (Dwan et al.

¹⁰ Based on the assumption that research findings are statistically normally distributed, the fact that 95% of published articles claim to have found statistically significant findings is a sign of publishing bias. This is the practice of not publishing negative or insignificant results. Publishing insignificant findings based on retesting experiments would have a corrective effect on the previously published significant findings. See also Ioannidis (2005).

2008), generates significant levels of false-positive findings, i.e. the claim that a new treatment or therapy works when in fact it does not. Ironically, this is exactly what the significance test aimed to prevent (Fisher 1925).

Significance tests are often combined with hypothesis testing as developed by statisticians J. Neyman and E. Pearson. In hypothesis testing, a researcher makes two assumptions about an experiment. The null hypothesis is that there is no effect, while the alternative hypothesis is often the opposite. The method also includes the study of effect sizes, which is the measurement of the strength of a relationship of measured variables. In this way, the method aims to be more rigorous and explicit about the outcomes of a statistical test than simple significance testing. Although Neyman and Pearson originally suggested their method as an alternative to Fisher's significance test, over time the two methods have merged to become the infamous 'Null Ritual' or null hypothesis significance testing (Berkson 1942).

Step one in the ritual is to define the null and alternative hypothesis without specifying what is expected to happen based on a researcher's contextual understanding of the research material. The second step is to apply the 5% cut-off rate for rejecting the null hypothesis and inferring statistical significance. If the experiment is statistically significant the researcher proceeds by accepting his research hypothesis. This procedure becomes ritualised by repeated application independent of context (Gigerenzer et al. 2004). This has led to intense debates (Cohen 1994; Goodman 1999a), where some have argued for discontinuing the teaching and application of significance tests (Schmidt 1996; Schmidt and Hunter 1997).

One of the key problems in the debate is the argument that the null ritual has led to a 'behaviour' rather than actual and contextually judicious, inferential statistical thinking (Goodman, 1999a). This refers to the sheer universal application of the conventional 5% cut-off rate as a signifier of seemingly actual and not just statistical relevance, but also the approach adopted by some journals (Gigerenzer et al. 2004) whereby papers not conducting such tests and finding statistically significant outcomes are not accepted for publication. Considering Neyman and Pearson's original motivation for proposing their method, that is to limit the number of mistakes made over many experiments, current practices of non-publication of insignificant findings or selective reporting can have very detrimental effects on the quality of actual evidence (Goodman 1999a) as we will see when considering such effects as identified from an EBM meta-analytic perspective.

Clinical research evidence is also generated in the form of observational studies and randomised controlled trials. There are various methodological differences between the two methods. Of particular relevance for the present

discussion are concerns over the level of control an investigator can exercise when conducting a study, potential bias that can arise due to a lack of such control and the implications this then has for drawing generalisable inferences. In a RCT, study subjects are allocated in a controlled randomised way to a study group and receive either the drug to be tested or, typically, a placebo. Once allocated, study subjects in both groups are treated in the exact same manner. Such controlled randomised treatment supposedly allows one to infer that any noticeable effect on the health of a study subject can be linked to the treatment itself.¹¹ This is thought possible because effective randomisation minimises selection and allocation bias, the former being differential treatment of study subjects when an investigator knows to which group they belong, while the latter means that any observable or unobservable causal effects will be equally distributed among the two study groups and, thus, minimise confounding.

To minimise these effects, where possible and ethical RCTs are performed as double-blind trials where neither investigator nor study subject are aware of which group the latter is allocated to. In their most encompassing versions, clinical trials are performed with up to 20,000 participants (Topol 2012). Such procedure is followed to ensure causal conclusions can be inferred, hence establishing the internal validity of the study. In contrast, external validity is concerned with whether the findings of a RCT are generalisable to a larger population of prospective patients. Sampling aims to ensure that the sample distribution is representative of the overall population and, thus, ensure external validity. However, as it transpires through the growing body of systematic reviews, this second line of EBM enquiry into the scientific robustness of medical evidence increasingly reveals some of the problematic foundations and processes whereby medical knowledge has been and still is generated and distributed.

Critical review of medical evidence

Ben Goldacre, a physician and EBM researcher, has framed these issues as the “broken information architecture of Medicine” (O’Reilly Media 2012). By that he specifically refers to his analysis which exposes the fact that there is a fundamental gap in the publishing of negative trial results (Goldacre 2012). Put differently, the structural bias towards publishing mostly positive trial results leads to an overstatement of the benefits of treatments (ibid.).

¹¹ While RCT’s prevent allocation and selection bias, they remain prone to other types of bias. For an overview and discussion see Jadad and Enkin (2007, 30).

To understand why this happens, it is useful to consider recent research by Ioannidis (2005), a leading meta-analytic medical researcher with an interest in the quality of medical research. In his study 'Contradicted and initially stronger effects in highly cited clinical research,' Ioannidis (2005) analysed actual medical publication patterns and how initial research findings were slowly corrected over time. For this, he analysed 49 of the most important published research findings¹² that were influential in developing

popularity for treatments such as the use of hormone-replacement therapy for menopausal women, vitamin E to reduce the risk of heart disease, coronary stents to ward off heart attacks, and daily low-dose aspirin to control blood pressure and prevent heart attacks and strokes [...]. Of the 49 articles, 45 claimed to have uncovered effective interventions. 34 of these claims had been retested, and 14 of these, or 41 percent, had been convincingly shown to be wrong or significantly exaggerated. If between a third and a half of the most acclaimed research in medicine was proving untrustworthy, the scope and impact of the problem were undeniable.

(Freedman 2010)

This problem may be explained to a large extent by selective reporting of research results and publication bias, as mentioned earlier. The former is the choice of data that scientists document, whereas the latter is the "tendency of scientists and scientific journals to prefer positive data over null results" (Lehrer 2010). A further problem is that trial samples seem to be overly ideal and, thus, for reasons such as young age and perfect single diagnosis, not representative of the larger population (Goldacre 2012, 177; Rothwell 2005). While in some cases this merely means an ineffective treatment for some patients, in others it has grave consequences by actually increasing morbidity (Pratt and Moyé 1995). Unfortunately, this does not prevent the healthcare system prescribing such treatments to the larger population.

Furthermore, when such a US Federal Drug Administration (FDA)-approved drug gets exposed via direct-to-consumer advertisement, rarer and potentially fatal side-effects may only then show up. Unfortunately, due to poor post-marketing surveillance mechanisms, it may take a long time before clear cause-and-effect relations between the drug and side-effects may be fully identified and proven (Topol 2012, 36).¹³ Because of how such evidence is

¹² This selection was based on the importance of journals and the numbers of citations (Ioannidis 2005).

¹³ This may be argued to be an effect of the sometimes non-representative character of sample selection with overly health participants in RCTs as well as the lack of systematic post-marketing surveillance mechanisms via doctors. See Goldacre (2012, 177) and Rothwell (2005).

generated in clinical trials via the evidence-based medicine approach, Topol, a physician, names it population-based medicine, while others have dubbed it 'mass medicalisation' (LeFanu 2011), though I agree with Rose that not much is gained from naming phenomena in such ways (Rose 2007).¹⁴ The key characteristic of population-based medicine can be further illustrated by looking at how the findings of a recent trial for a cholesterol-lowering drug have been presented (Topol 2012; Goldacre 2012) and applied. The outcome was a reduction of events from 4% in the placebo group to 2% in the actual drug-taking group. Typically, such an outcome is deemed clinically significant, as it reduces the relative likelihood of the probable pathological event by 50%, which is also how the effectiveness of the novel drug is marketed. The problem, however, remains that it is impossible to clearly identify which two patients out of the 100 treated will benefit, thus leaving 98 drug-taking patients out of 100 without such benefits while having to take the prescription (Topol 2012).

Surveillance-medicine

The logic of population-based medicine is not only underlying actual treatment of patients with diseases but also the area of prevention or surveillance medicine as some have named it (Armstrong 1995). A contemporary example to illustrate the operative mechanisms of surveillance medicine is cancer screening. In the USA, men over fifty are screened for prostate cancer (Ablyn 2010). The screening is typically performed with a PSA test that measures the levels of prostate specific antigens in the blood.¹⁵ Unfortunately, the test results in a relatively high number of false-positive findings, leaving 250,000 men per year having one or multiple biopsies before they find out that the initial screening was not an accurate indicator of actual prostate cancer (Topol 2012). Despite the fact that over 15% of men going through screening are diagnosed with prostate cancer, only 3% die from it, which indicates that, to a very significant degree, prostate cancer is in fact nonaggressive in its prevalence. Nevertheless, all those diagnosed with prostate cancer typically undergo surgery followed by radiation treatment. Men, in these cases, suffer

¹⁴ This is because the term 'medicalisation' obscures the intricate details in which changes in medical epistemology and clinical practice affect different people with different health and pathological states in different regions and, thus, may be said to be oversimplifying.

¹⁵ PSA is a protein produced by the prostate gland (<http://www.cancer.gov/cancertopics/factsheet/detection/PSA>- ac. 02.08.2014).

overdiagnosis and, in undergoing standard treatment, also experience and suffer various kinds of unnecessary pain and debilitating side-effects.

In fact, the inventor of the PSA-test, which has become the standard method for screening male populations and assessing the likelihood of prostate cancer, has recently declared it 'The Great Prostate Mistake' in an open letter to *The New York Times* (Ablin 2010). In the letter, he explains the scale and problem of the misuse of the test as well as its original intention. Instead, he argues for a specific and case-by-case application. The UK National Screening Committee, a government body assessing evidence for screening programmes and making policy recommendations, has recently voted against a screening programme based on the PSA test¹⁶ mainly because of the risk of overdiagnosis and poor effectiveness in preventing death from prostate cancer (Sandblom et al. 2010).

To some degree, then, it appears that the argument about variations in clinical practice stemming from physician uncertainty due to what was essentially seen as information overload may, at best, have only been part of the explanation. More importantly, it seems to have obscured much larger distortions in the underlying processes of medical knowledge production and distribution. The cases of cancer screening seem to provide some evidence as to the degree of imprecision in clinical testing and the resulting space for expert medical assessment and, thus, variation by physicians. In contrast, medical evidence following an EBM methodological approach builds on a set of seemingly axiomatic stratagems which aim to establish more reliable, better controllable and, as we will see, ever finer diagnostic mechanisms for identifying, forecasting, delineating and explaining various degrees of actual and anticipatory pathological states in bodies or parts thereof.

Following the medico-statistical mechanisms analysed in more detail in the sections to follow, notions of normalcy and pathology in EBM have internalised a logic resembling the calculative, flat rationality of game theory approaches.¹⁷ Thereby, and as materialised in the form of decision-aids, they map out a territory of preferable action, non-action and probable and potentially grave bodily consequences. When following the evidence generated, the above is aimed at panning out at population levels of some kind. Simultaneously, however, due to its probabilistic nature it leaves significant actual variation in individual health outcomes. Among further aspects, of key concern for this research is the amalgam of probabilistically-

¹⁶ "Experts scrap prostate screening proposal," *BBC News Health*, December 6, 2010, <http://www.bbc.co.uk/news/health-11930979>; ac. 01.03. 2013.

¹⁷ For the shortcomings of Decision Analysis see Bursztajn et al. (1990).

determined pathological states inferred from population samples-based medico-statistics, specifically when used for individual decision-making.

Data and statistics

The history of data and statistics is witness to the sources of interest driving the collection of data, the underlying theories based on which the statistical insights were viewed and argued about as well as claims of authority over the epistemological and interpretative frame and action on them. Each of these will be explored in the following sections.

Early motivations for data collection and statistics revolved around issues of interest to the state and the affirmation of its power. Census techniques were applied for example in 1548 by the Spanish to take stock of Peru.¹⁸ William Petty, an early influential British figure in statistics, for example, had directed the stock taking of Ireland's territory, buildings, people and livestock, which then also facilitated its exploitation by the English in 1679 (Hacking 1990, 17). Ideas for the institutionalisation of census techniques were debated in the late 17th century when Leibniz, among others, suggested that, for the establishment of a Prussian state, a central statistical office should be set-up in order to know a state's "true measure of power," i.e. the size of its population (ibid., 18). This measure was to be achieved based on his proposed 56 categories for evaluation, including "number of people by sex, number of able-bodied men who might bear weapons, number of marriageable women, child mortality, life expectancy, distribution of diseases and causes of death (ibid.,19).

Apart from state interests, early statistics was driven by commercial interests and the pricing of annuities, the then common way to raise capital for states and cities. In principal, the challenge was to accurately estimate life expectancies so as to balance in-payments and future out-payments. This undertaking was severely challenged by the lack of sound mortality statistics. The unreliability of existing mortality statistics can be partly explained by the curiosities of the mechanisms for collecting such data. For example, before 1830 parish churches of England were responsible for recording baptisms, marriages and burials, which were the source for generating life tables and calculating mortality rates. Due to the inherently religious perspective, the records noted baptisms rather than births, and burials rather than deaths. Thus, dissenters, who were prevalent at the time, the religiously indifferent and in some cases even poor Anglicans may not have been reflected in the

¹⁸ There are also various earlier examples as documented by Harold Innis (1948).

records (Eyler 1979, 37). Customs in early 19th century England demanded some kind of celebration at baptism and registering a baptism incurred a fee, both of which may have discouraged complete registration. Nevertheless, parish records played an important role as they were held as evidence for tracing ancestry and assuring inheritance. This deficiency was eventually corrected by the introduction of the Registration Act in 1836 and the General Register Office (GRO) which was responsible for civil registrations of births, marriages and deaths. The GRO today is part of the Office for National Statistics. As we will see further down in the case of trials of novel pharmaceutical products, somewhat similar problems with data persist as a consequence of incentive structures, albeit in the case of clinical research in more refined ways and due to issues such as selective reporting and publication bias.

Much of the early developments were driven by a fascination with statistics' seeming ability to bring order and regularity to complex events, such as the duration of human life, crime and suicide rates. In this regard, in 1832 Charles Babbage, a British mathematician and philosopher, proposed a broad plan to measure "the constant qualities belonging to our solar system" (Hacking 1990, 57). In his letter to the Scottish physician Brewster, he proposed, among other things, to measure the quantity of air consumed per hour, the proportion of sickness among the working classes and the power of steam engines in Cornwall (*ibid.*, 58). Constants for him were "[A]ll those facts which can be expressed by numbers in the various sciences and arts" (*ibid.*, 55). Thus, according to Hacking, constants for Babbage were more akin to rules or regularities, rather than fundamental laws such as those in physics as we know them today. In the spirit of that time, many laws were unsurprisingly identified and even more attempts to find them were almost obsessively undertaken, including whether one's "ability to cultivate flowers or sing cheerful songs were useful indices of respectability and morality" (*ibid.*, 29).

Concerning claims to authority, a relevant example to consider is the case of suicide statistics. In France, philosopher Marquis de Condorcet was an early eminent figure in statistics. For him, moral sciences were to be understood as "all those sciences that have as their object either the human mind itself or the relations of men one to another" (*ibid.*, 38). Most early statistical data was confined to vital statistics (births, marriages and deaths) which were less interesting to him as they did not relate to individual behaviour. Nevertheless, there was a connection between the original census-driven developments of statistics and those of interest to Condorcet and his contemporaries. That was an interest in 'deviant' behaviour, particularly in subgroups, such as suicidal French men, who did not contribute to the growth

of France.

Importantly, such early behavioural data was approached with a firm belief that laws of society exist in the same way that laws or 'constants' of nature exist. It is important to note that statistical data did not operate in a void but was and still is closely intertwined with existing theories of what such data should mean and how it fits in existing classification schemes, for example in medical nosology¹⁹ and criminal behaviour. Esquirol, a French medical researcher, became a leading figure in the argument that suicide is a medical issue rather than a moral one. He essentially argued that treatment of madness is the responsibility of medicine and suicide is a kind of madness, hence within the realm of physicians rather than the moral sciences (ibid., 65).

Furthermore, according to the organic theory of disease popular at the time, all diseases have their origins in defective tissue or organs. For the French this naturally led to the need to dissect bodies of suicide deaths. However, they were not able to identify substantial differences between suicidal and non-suicidal brains (ibid., 70). From the perspective of this research and following the logic of this example, it is useful to note that such questions of disciplinary responsibility and control over the epistemological and interpretative frame also dictated potential actions following the collection of such data. As will be discussed in the context of anticipatory surveillance medicine practices, this logic had a central role in determining the meaning of normalcy and pathology.

Another concept that has come to play a key role in the success story of statistics to this day is that of the 'average man.' The Belgian astronomer and advocate of statistical thinking and practice, Adolphe Quetelet, was not satisfied with detecting mere regularities in deviant behaviour and crime rates. He was interested in finding the distinct laws that govern people's behaviour like those known in astronomy. In 1844 he announced that many human characteristics and behaviours are normally distributed among the population (ibid., 105).²⁰ This is the same well-known distribution as when one tosses a coin many times. The arithmetical average of a population, such as for example the average family having 2.3 children, could not necessarily be considered a meaningful quantity in itself. However, Quetelet framed the attributes of a statistically 'average man' within the boundaries of a racial type. That is, instead of describing populations with reference to extrinsic characteristics, such as language, their geographic locality or their ruler, which was custom at the time, Quetelet established an account of a race by means of their objectively measurable physical and moral, thus 'intrinsic,' characteristics,

¹⁹ The classification of diseases in medical science.

²⁰ This is also known as 'the bell-shape distribution.'

such as men's height but also 'knowable' moral traits (ibid., 107). Retrospectively and significantly, what happened at that moment, according to Hacking (ibid., 108), was the introduction of a "new kind of information about populations and a new conception of how to control them," as, for example, via the introduction of social policies to retain or, more perversely, alter the measurable average qualities of a population of race. As history sadly witnessed, eugenics, the idea to improve the human race by controlled breeding of desirable and heritable characteristics, was not long to follow.

Indeed, this was also very much the attitude of the highly active reformer William Farr, a qualified doctor who became the first compiler of scientific abstracts at the newly formed General Register Office for England and Wales. He was a fervent proponent of statistics and coupled it with his background in medicine by producing an ever-growing body of statistics for sickness and corresponding laws of sickness. As indicated above, Farr was specifically convinced of the deterministic power of statistical laws with regards to human populations. He did not subscribe to statistical fatalism, that is to say the doctrine that, if a statistical law applies to a group of people, it will necessarily constrain the freedom of individuals in that group. Unsurprisingly, though, given the utilitarian spirit of Farr's time, he "maintained a strict social determinism: The members of the governed class remained bound by statistical law, albeit one that was chosen by a well-meaning bureaucracy" (ibid., 118).

Before continuing the discussion of Quetelet's statistical 'average man' and its repercussions, it is worthwhile briefly exploring a bit further the deterministic understanding of statistical regularities at the time. How were such regularities accounted for? The conundrum Quetelet, Farr and their contemporaries struggled with was that, if statistical laws were to hold, then individuals were not free to choose. This was problematic, particularly for a western liberal mentality, as it was difficult to imagine that an individual should not be free to choose whether or not to commit suicide. In Paris in 1813, the 243 drownings in the Seine (ibid., 66) during that year were perceived to be mainly of a voluntary nature. To resolve this conundrum, a theory at the time was that, while individuals were free to choose, such choices were thought of as minuscule causes which cancel each other out in a large number of individuals and allow statistical laws to surface. As Quetelet had it: "The larger the number of individuals, the more the individual fades out and allows the series of general facts to predominate, the facts which depend on general causes and in virtue of which society exists and is conserved" (Quetelet [1832, 81] quoted in Hacking 1990, 123).

In fact, Quetelet's conception was unintentionally substantiated by

Poisson's mathematical proof of the law of large numbers, which states that, over a large number of trials, the mean of the trial results will stabilise. Even though Poisson suggested it as a mathematical theorem, it became accepted as an a priori law of social behaviour and other phenomena (Hacking 1990). But the stability of statistical findings over time and, thus, their usefulness as a source of information for political economy aiming to regulate medium- to long-term social and economic conduct was questioned (Porter 1988, 153).

Arguments were voiced on either side, yet with little data to sustain them for long. Firstly, Jevons, an economist critical of mechanistic deterministic laws or pure chance, pointed out early on that slight changes in initial conditions and applications of the same 'laws' could yield very different effects (ibid., 177).²¹ In this light, for Jevons statistical 'laws' were not universal but, rather, context-dependent. Secondly, statistics, with its focus on averages at the time, was essentially also seen as a strategy for ignoring individuals and their perturbations, simply averaging them out (ibid., 152). Such a procedure was deemed uninformative from a scientific perspective, as no new knowledge could be gained from it. According to d'Amador, a Spanish physician and early proponent of homeopathy, "to resort to probability is to appeal to chance and to give up the possibility of certitude" (ibid., 159). Lastly, for Bernard, a French physician, positivist and ardent researcher of physiology, the numerical methods of statistics and mathematics were not the problem; rather, the complexity of vital phenomena remained ill-understood and, thus, was no fertile ground for inferring useful information following such numerical methods. Bernard proposed experimental physiology to counter these shortcomings rather than averaging out errors (ibid., 160). The applicability of averages in medicine was seen as immoral, as successful treatment of some based on such averages comes at the cost of death for those unaccounted for therein (ibid., 157). From the perspective of this research, these critiques are significant and remain relevant concerning the role of probabilistic evidence as a useful source for decision-making by individuals within the context of cancer screening.

Despite these critical voices, the law of large numbers was the dominant understanding around the nature and power of statistical laws at the time. Significantly, this was coupled with a strong conceptual interpretative perspective, both in the case of medical knowledge about the human body and the causes of diseases as well as social processes and society more broadly. Broussais, a French doctor and outspoken promoter of the organic theory of disease, also introduced and championed a novel conception of healthy and pathological states of the body. Before him, those states were

²¹ This is now known as 'the butterfly effect' in chaos theory.

considered separate in nature, so that knowledge about one could not be applied to the other or indeed effect a transgression towards the other. Broussais, however, argued, that “phenomena of disease are the same in kind as those of health, from which they differed only in intensity” (Comte [1851,1] quoted in Hacking 1990, 160).

As we will see further below, this logic became the principle upon which epidemiology and evidence-based medicine operate and, thus, carries major implications for the ways in which a normal state of health and deviations from it are considered today, both in medicine and society at large. Broussais argued that diseases are nothing else but a local irritation due to a change in intensity of excitation of tissues. Physiological medicine was, thus, responsible to establish how “excitation can deviate from the normal state and constitute an abnormal or diseased state” (Broussais [1828, 63] quoted in Hacking 1990, 82). His understanding that such deviations from a healthy state were of a continuous character inversely meant that appropriate treatment would smoothly regulate the irritated tissue back to a normal, hence healthy, state.

It was precisely the combination of a conceptual characterisation of a ‘normal’ and ‘pathological’ state in combination with Quetelet’s notion of the ‘average man’ which was appealing to Comte, the French philosopher and founder of sociology. Importantly, Comte wholeheartedly subscribed to Broussais’ conceptualisation of the pathological as a deviation from the normal state and extended it to the study of society. The normal state for Comte was what we should strive for (Hacking 1990, 168) and its widespread adoption and use by statisticians has stayed with us. This can easily be demonstrated, for example, when quickly browsing the ‘Health Survey for England 2011’ as commented on by *The Guardian* newspaper datablog.²² The article starts with the assertion that “[T]he average adult in England is overweight, as measured by mean body mass index (BMI),²³ and obesity rose among England’s children in 2011” (Burn-Murdoch 2012). The three characteristics outlined above all manifest in this quote. A sample of a population was taken and the mean of an intrinsic characteristic calculated. Based on the continuous scale of the BMI, the mean deviates from what is considered ‘normal’ and is, thus, postulated as pathological, in this case overweight and obese.

Ironically, the BMI was invented by Quetelet himself in 1832 and published as part of his seminal work *A Treatise on Man and the Development*

²² Burn-Murdoch, J. 2012. “Health Survey for England: the key data on obesity,” *The Guardian*, December 20, 2012, <http://www.guardian.co.uk/news/datablog/2012/dec/20/health-survey-england-obesity-trends-data>; ac 13.01. 2013.

²³ The BMI does not measure actual body fat but is a heuristic proxy for estimating body fat based on an individual’s height and weight.

of His Faculties (1832), formerly known as the Quetelet index and renamed as the body-mass index by Keys et al. (1972). The cut-off rates²⁴ are calculated with reference to increasing risks of diabetes and cardiovascular diseases. However, when Keys suggested its superiority compared to other weight/height ratios, he meant it to be used for population studies only because, due to its heuristic character, it is an unsuitable measure for individual diagnosis (*ibid.*). As evidence has indicated in the meantime, cut-off rates do not work universally across different populations (WHO 2004) and the BMI itself performs poorly in predicting cancer (Romero-Corral et al. 2006). The problem of average values following population studies and the temptation to meaningfully use them for supporting decision-making by individuals constitutes a key focus for this research.

To summarise then, the history of data and statistics exhibits a number of key developments which may help us understand the potential implications it has on the production of medical knowledge as well as critically assessing more recent data-based approaches to medicine. The key developments for our purposes are that early statistics was motivated to manage populations and risk, and that data collection and analysis is always informed by an interpretational framework (theory) that dictates what such data means and how it should be classified.²⁵ Furthermore, statistical laws were heralded as biologically and socially inclining if not deterministically understood, while the introduction of the law of large numbers further substantiated their reliability. The notion of the 'average man' was introduced as archetypal of a racial population and, thus, meaningful information. Lastly, the 'normal' and the 'pathological' were defined as different medical states on a continuum, their difference being one of intensity and not quality.

History of the probabilistic approach in evidence-based medicine and the notion of the 'normal'

In this section, I will attempt to trace a genealogical thread of the specific role quantification of physiological parameters and the subsequent calculation of their statistical probable distribution among populations have come to play in the production and application of medical knowledge. I will discuss historical struggles for more encompassing etiological conceptions relevant for medical knowledge and the specific capacity that statistical analysis and insights

²⁴ Cut-off rates for what is considered normal weight, overweight and obese are set by the World Health Organisation (WHO).

²⁵ Of course, there is the potential that data and theory may be informing each other over time.

should have furnished. I will argue that, while quantification and statistical analysis were eventually embraced more broadly in medical research with the introduction of the notion of 'risk factors,' this led to subsequent changes in the conceptual understanding and definition of the 'normal' and the 'pathological.'

The spectrum of probabilistic risk factors has specifically proven to be influential when, through the promotion of health and preventive medicine, responsibilities for one's health slowly extended surveillance modes by government to self-surveilling individuals. As I will argue, an intricate and problematic link has been established between how probabilities of being affected by a disease are conceptualised by the healthcare industry and an individual's assumed role in reducing risk factors to prevent illness and retain health. This critical analysis will generate the necessary ground on which I will reflect on the potential and problems of Big Data in relation to medical knowledge in this chapter as well as exploratory search and individual decision-making in the next chapter.

As discussed above, one of the early proponents of compiling and using vital statistics in the context of government policy was William Farr. At his time, and as a result of the consequences of the industrial revolution, the living and working conditions for the urban working population were terrible as evidenced in high urban mortality rates (Hammond and Hammond 1917). These circumstances also gave rise to concerns about social unrest and the threat of revolution, which the governing class felt the urge to do something about (Susser 2009).

This was a topic of interest to Farr throughout his long and active career in promoting vital statistics and specifically mortality rates as a gauge of quality of life. Among his early work is a paper entitled "On a Method of Determining the Danger and the Duration of Diseases at Every Period of Their Progress" (1837). In his first letter for the Annual Report of the Registrar General in 1839, he argued that, while advances in medical knowledge remain important, much greater influence on public health can be achieved by registering causes of deaths (Farr [25] cited in Eyler 1979, 110). He claimed that it is less difficult to prevent diseases than cure them and, for this reason, it is necessary to understand what causes them. The role he saw for himself at the G.R.O was to

show the agency of these causes by numerical facts, and measure the intensity of their influence. In exhibiting the high mortality, the diseases by which it is occasioned, and the exciting causes of disease, the abstract of the registers will prove, that [...] a considerable proportion of the sickness and deaths may be suppressed by the general adoption of hygienic measures, which are in actual but partial operation.

(Farr [26] quoted in Eyler 1979, 124)

With this, Farr was proposing the use of statistical methods not only to count the dead categorised by cause of death, but also, and more importantly, on this basis to be able to calculate and project the probable influences of the causes of a disease. As indicated earlier, Farr saw himself not merely as a compiler of statistical facts, but equally as a reformer. In this sense, his proposal can also be read as a much broader definition of medicine which should include epidemiological methods and operate as preventive medicine. His proposal for such a preventive medicine, however, only included suggestions for housing and sanitary reform, including street cleaning, better ventilation by widening streets as well as vaccination and individual hygiene (ibid., 127).

Furthermore, it is important to highlight his influence in creating nosologies, which came to be used by the G.R.O and beyond. Diseases in medical science are classified in a nosology. There are different nosologies and diseases can be classified according to their cause (etiology), the mechanisms by which diseases are caused (pathogenesis) or the symptoms a patient experiences. The nosology Farr developed was important because this categorisation consolidated the medical theories of death and, thereby, dictated the way in which deaths were counted nationally. By way of the magnitude of the numbers counted, his nosology affected the understanding and interpretation of causes of mortality at a national level.

Nevertheless, his nosology was not without its critics, primarily public health officers with a deeper medical interest and a belief that the identified variations in local mortality rates were not well explained by Farr's zymotic disease category comprising the epidemic, endemic and contagious diseases.²⁶ It was felt that the zymotic class contained diseases too diverse in character and that, by using more fine-grained disease categories, greater insight could be gained. A more detailed examination indicated that major causes for variations in local mortality were particularly due to pulmonary diseases as well as dysentery, diarrhoea and cholera. Under John Simon, the successor of Edwin Chadwick as the public health administrator in office at the time, this led to the monitoring of variations and investigation of rapid changes in local mortality rates due to such causes.

This is a very relevant moment of transition in the use of health statistics. Whereas previously it was customary to assess the health of the population

²⁶ Such as smallpox, cholera, typhus, plague and influenza as well as other diseases, such as cow pox, glanders, hydrophobia, syphilis, erysipelas, puerperal fever, measles, scarlet fever, whooping cough, dysentery and diarrhoea (Eyler 1979, 82).

relatively broadly by way of general mortality statistics, under Simon the same methods were employed to identify, analyse and explain specific disease developments. While from a narrow physiological-medical perspective Farr's zymotic disease can be viewed too broad a category for precise medico-pathological counting, it could equally be seen as the cumulative casualties resulting from the deplorable living and working conditions of the working class. That seemed to be his intention, as he explained in his first letter to the A.R.R.G in 1839 by relating to this class of diseases as the "index of salubrity" which in his opinion could be controlled and prevented (Farr [26], quoted in Eyler 1979).

Similarly, a group of advocates struggled for the development of a social medicine in Germany. Virchow, a German physician and politician, published in cooperation with a group of like-minded peers the journal *Die Medicinische Reform* where they argued for a reconsideration of the field of medicine as a social science and the urgent need for medical reform. The principles upon which they based their proposition for reform were that the health of people is a matter of direct social concern, that social and economic conditions have an important effect on health and disease, and that these relations must be subjected to scientific investigation. As a logical consequence steps must be taken to promote health and to combat disease, and that the measures involved in such action must be social as well as medical (Rosen 1974, 67). Their proposal was submitted as a draft for a Public Health Law in 1849, but was unsuccessful and the hopes for a health reformation died with it.

Nevertheless, the significant statistical differences in life expectancy between workers and the upper classes derived from those statistical studies by Farr and others²⁷ provided strong arguments for improvements via reform. It is interesting to note here, though, that the reform measures suggested and implemented were mostly infrastructural and did not include aspects concerning social and economic conditions of labouring people at the time. Some of these concerns were only later taken up as issues of 'occupational health' as a result of labour movements in the second half of the 19th century. The broader social concerns in relation to health which were identified by Villermé, Virchow and Wakley, however, were of little wider political interest for a long time.

In England, and during Chadwick's time as commissioner of the General Board of Health,²⁸ sanitary reform was based on indices, such as mean age or general mortality rates, and focused on infrastructural and environmental improvements, such as sewage systems and access to fresh and clean

²⁷ Such as Villermé, a reformist French physician (Hacking 1990, 74).

²⁸ The General Board of Health was abolished in 1854 (Hacking 1990).

drinking water. In contrast, Simon, in his capacity as Chief Medical Officer, initiated a policy of preventive medicine which was based on death rates from specific causes (Eyler 1979, 137). With the presentation of his *Sanitary Papers* report in June 1858, he aimed to initiate a second phase of sanitary government. For him, a novel approach was required to advance public health, which differed from previous efforts by Chadwick that he saw as based on crude generalisations and preoccupied with engineering. Instead, the novel approach should become focused on preventive measures and based on scientific knowledge (Lambert 1963, 267). Influential in the development of epidemiological orientation was germ theory. More specifically, the publication by R.E. Koch, a German physician and founding figure of microbiology, on the tubercle bacillus²⁹ in 1882 demonstrated the ability of the approach to identify a microorganism responsible for causing tuberculosis. Epidemiological orientation, thus, shifted from environmental causes, popular during sanitary reform, towards bacteriological aspects and transmission patterns of diseases.³⁰

Early cases where quantitative measurements and simple statistical analysis were deployed in the context of medical treatments were the vaccination for small pox³¹ and bloodletting.³² Bloodletting as a therapy to cure or prevent illnesses has been practiced since antiquity and was based on the medical concept that, broadly speaking, health depends on the balance of bodily fluids. It is a therapy which was also vigorously applied by some French doctors, among them Broussais. This is unsurprising, given his novel conception of normal and pathological states as different states of the same kind. Bloodletting, thus, must have seemed to him as logically adequate to rebalance a patient's pathological state towards a healthy state (1832).

Louis, another French physician, collated 100 cases of pneumonia treated with bloodletting and compared them to a control group. This evidence, he asserted, proved the uncertain effectiveness of the therapy.

²⁹ A small lesion in the lungs, characteristic of tuberculosis (Pschyrembel 2010).

³⁰ It was not until the 1980's, however, that social epidemiology re-pronounced the importance of considering the implications of the social environment onto the health of individuals and society (Honjo 2004; Krieger 2001).

³¹ in 1721, Rev Cotton Mather together with Dr. Boylston compared mortality statistics for small pox based on natural causes with those vaccinated, finding that, in the former case, one out of six proved fatal, while only one in ninety-one of the inoculated cases were fatal (Shryock 1961).

³² It has to be noted that there is a long history of earlier more or less successful such trials to quantify, dating back to the time of Hippocrates counting days of critical points of fevers as well as the now extinct area of iatrophysics which, roughly from 1600 to 1750, aimed prematurely at quantifying medicine (Shryock 1961).

While some found the evidence credible, others were critical of it and pointed to two important shortcomings. Firstly, Louis left unclear what level of quantitative evidence would qualify as a significant finding and, thus, rule out significant influence of confounding effects. Secondly, while statistical averages were interesting, such knowledge would not be applicable for effective therapy of individuals. As a result, the concept of “l’homme moyen would lead to indiscriminate, routine treatments” (Shryock 1961, 19). As we now know, it was not until 1925 with the introduction of Fisher’s statistical significance test that statistical rigour was strengthened. ‘Indiscriminate routine treatments’ however, is precisely what seems now to be everyday clinical practice as highlighted earlier with regards to population-based medicine.

Risk-factor epidemiology

Apart from this early example above, quantification of medicine had an uneasy reception early in the 19th century. This was exemplified by critical voices being raised by influential medical professionals. Auber, a French physician, published his *Traité de Philosophie Médicale* in 1839 where he argued that “many physiologic and pathologic phenomena were unmeasurable, since the process could give numbers but not the qualities of things” (Shryock 1961, 224). This is an important criticism, which is also reflected in the earlier discussion of the difficulties in applying BMI cut-off rates across diverse populations. The problem at the heart of this seems to be that, while measurements of certain aspects, such as glucose-levels in a person’s urine or blood in the case of diabetes or PSA-levels in relation to prostate cancer, are given as cardinal numbers on a continuum, they cannot in themselves necessarily speak of the qualitative pathological changes occurring in a person. While this general observation is hardly surprising, it does warrant additional attention, particularly in relation to the setting of thresholds to demarcate clinical normalcy from pathology, to which I will turn now.

It was not until after the Second World War that Archibald Cochrane, among others, proposed to frame the epidemiological focus on integrating environmental, social and microbiological aspects in order to understand and identify causal factors for diseases. Prevalent at the time were significant increases in chronic and non-communicable diseases, such as cancer and cardiovascular diseases. Thus, from a medical perspective, the focus of epidemiology became more specific again, but likewise changed its underlying etiological paradigm. Rather than a single spreading causal agent, as in the case of miasma- and germ-theories, chronic disease epidemiology

followed a 'risk factor' and 'web of causation' paradigm. According to Susser and Stein

Under this paradigm, the logical approach for epidemiologists is to seek to identify risk factors - exposures or characteristics that confer increased risk - for disease, rather than to look for a one-to-one relationship between cause and disease. The logical approach for public health intervention is to alter the risk profile of individuals within the population.

(Susser and Stein 2009, 12)

Risk-factor epidemiology operates with observational studies, such as case-control and (prospective) cohort-studies. In contrast to RCTs discussed earlier, sample selection is performed in relation to hypothesised causal attributes and exposure to these causal attributes rather than being perfectly randomised. The goal lies in the isolation of these causal attributes and risk factors. Evidence from the method is deemed less reliable due to the various possibilities of bias which are accounted for with RCTs as previously discussed. Also, observational studies are less interested in providing insights as to the distribution of phenomena in demographic terms, that is to say whether all who exhibit a certain risk factor are necessarily also at risk of the related disease.

Once smaller observational studies provide indications for specific risk factors, such evidence can then be followed up with larger-scale studies to provide stronger, statistically-significant evidence. This was, for example, the case in the much-cited research which related smoking to lung cancer (Doll and Hill 1950, 1954) as well as the Framingham Heart studies. The Framingham Heart studies are ongoing longitudinal prospective cohort studies concerned with understanding risk factors in relation to the development of heart diseases. It is set up as the study of residents of the town Framingham in the USA and cited for having 'established' evidence relating smoking, elevated levels of cholesterol and obesity with increased risk of heart diseases (ibid.). Participants undergo a detailed medical examination at regular intervals. The research study thereby allows the tracking of the emergence of pathological incidences over time and statistical determination and prediction of the correlation between physiologically measured parameters (such as blood pressure, cholesterol and level of obesity) and future illnesses, establishing in this way the former as risk factors to potentially prevent.

The case of diagnosing various forms and particular stages of diabetes, a metabolic disorder identified for about 3500 years (Greene 2007, 84), reveals the issues with quantifying physiological parameters and the definition of thresholds of normal and pathological states in more detail. Before the

discovery of insulin as an effective treatment of the disease in the 1920's, clinical focus in the case of diabetes, as well as more broadly in medicine, was on treatment of patients with symptoms. The introduction of insulin to therapeutic practice had significant positive effects on the life expectancy of patients with diabetes, particularly children.

However, as people with diabetes lived longer, a number of hitherto unknown diabetic conditions, such as diabetic eye disease that might result in severe vision loss or blindness, and other conditions surfaced.³³ Subsequent medical and pharmacological research focused on developing diagnostic tools to support identification of diabetic patients unaware of their pathological state as well as an oral and more convenient form of therapy than was available with the hypodermic needle for injecting insulin. Diagnostic urine test kits were developed and increasingly spread via GP practices, gymnasia, summer camps (ibid., 98), thereby starting establishing a surveillance infrastructure in the USA. This helped to identify what at the time was popularised as 'the million hidden diabetics,' that is to say people with mild forms of diabetes very often treated with diet and exercise.

Nonetheless, urine-based testing had a number of disadvantages compared to blood-glucose level testing. The most important of them was that it would only indicate sugar levels above a relatively high threshold³⁴ and provide only relatively rough indications of sugar levels, which may not necessarily be indicative of blood-sugar levels. Thus, it would not allow for more fine-grained distinctions of various intensities of diabetes and corresponding therapy. The increasing use of blood glucose level based screening and diagnostic methods changed this by providing a continuous and increasingly sensitive scale of blood glucose levels in tested people.

Simultaneously, it led to rising numbers of people with 'borderline test results,' that is to say results near the set level for diagnosing overt diabetes. In relation to family histories with diabetes occurring, these borderline cases were subsequently considered as at risk for diabetes. While the terminology for classifying this risk group varied from protodiabetic, chemical diabetic, latent diabetic, stress diabetic and pre-diabetic, they shared a common characteristic, namely the lack of symptoms. After some debate, these various typifications were consolidated and reclassified for clinical diagnosis as chemical, mild or early diabetics in the early 1960's (Greene 2007, 100, 106). Consequently, numerous people previously considered healthy were now

³³ Such as the "susceptibility to infections, poor wound healing and vascular diseases," not to mention social stigma (Greene 2007, 87).

³⁴ The renal threshold of urine-sugar level is regulated by the kidney and varies from person to person.

thought to be ill and, thus, treated. In turn, the category of pre-diabetes was reconceptualised as people who develop symptomatic diabetes only under bodily stress, such as during pregnancy. Although it was found that such symptomatic diabetes typically recedes after pregnancy, other studies suggested significant statistical correlation with overt diabetes in later stages of life as compared to non-pregnant women. Such pre-diabetes “was recast as an early warning sign of an underlying, incipient diabetes” (ibid., 106).

To summarise, there was a shift from seeing specific diseases as incurable towards reframing them as chronic diseases which could and, from a public health perspective, should be predicted, surveilled and continuously treated. Surveillance was performed by the regular measurement of physiological parameters considered indicative of future pathology. What remained unchanged from earlier epidemiological approaches, however, was the methodology for calculating the influence of risk factors or treatments in relation to broad sociodemographic and increasingly specific genetic factors (Rose 2006). This methodology, termed ‘population thinking’ in epidemiological jargon, has become one of the foundations of clinical research, clinical epidemiology and, thus, evidence-based medicine to this day.

The conception of a set of risk factors responsible for causing a disease is also reflected in the transition of larger medical paradigms from what has been termed ‘hospital medicine,’ dominant in the 19th century, towards a preventive and surveillance medicine in the early 20th century. Hospital medicine had as its object of study the gross anatomical structures of organs and tissues that are visible to the eye. Distancing themselves from the ill man, medical investigators reduced symptoms to secondary indicators rather than defining features of disease. Diagnosis, in turn, was established through physical examination of observable structures. Pathological anatomy was technically supported by the invention of a number of ‘scopes,’ such as the stethoscope. Importantly, whereas abdominal pain, for example, was previously seen as the illness itself, in hospital medicine this was merely considered to be a symptom which, through examination by an experienced physician, might have been linked to a sign, such as abdominal tenderness. Based on the prevailing pathology, both symptom and sign would then be used to infer a hidden pathological lesion. As Armstrong (1995, 402) succinctly sums it up, whereas “pathology in hospital medicine had been a concrete lesion, in surveillance medicine illness it is in perpetual becoming.”

Risk-factor epidemiology can, thus, be said to have fully established the mechanisms for population-based and surveillance medicine, thereby closing the circle by embracing Broussais’ much earlier conception of health and pathology as being merely different states on a linear and measurable

gradient. In surveillance medicine, the notion of a risk factor came to problematise the 'normal' and "encompasses any state or event from which the probability of illness can be calculated" (ibid., 401). It is, indeed, the early identification of probable risk factors which appears increasingly and specifically problematic. As Greene (2007, 112), a historian of science and doctor of medicine, aptly comments in relation to the interpretation of pre-diabetic test results, "[By] equating the linear gradient of physiological parameters with the temporal progression of disease, the concept of pre-diabetes invested borderline test results with a sense of pathophysiological urgency."

Reflections on concepts of normality and pathology

For Georges Canguilhem, French philosopher and historian of science, the question of what constitutes normal and pathological states is at the heart of the conceptual and operative understanding of medicine (1994). I will very briefly summarise the different conceptions and developments of normal and pathological states that have been discussed above and provide diagrams to visualise the changes observed. Loosely, three phases will be differentiated: the pre-Broussais phase, the Broussais phase and the risk factor phase. Such a diagrammatic understanding will be useful when evaluating more recent data-driven approaches to healthcare and assessing the kinds of contributions they might be able to make.

Pre-Broussais phase

During this phase it was held that physiology and pathology were two distinct domains of phenomena and knowledge. Consequently, it was thought that having knowledge about one of them could not inform knowledge of the other. Signs and symptoms were indicators of pathology.

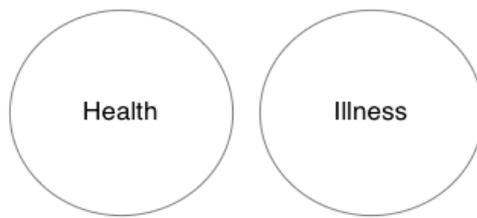


Figure 2.2: Relation of health and illness in pre-Broussais phase

Broussais phase

Broussais conceptualised the normal and pathological as different states on a gradient of the same kind. The difference between the two states, according to Broussais, can be explained by changes in levels of intensity. Pathology, thus, was seen as an intensification of normal states.

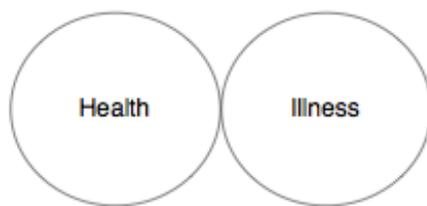


Figure 2.3: Relation of health and illness in Broussais phase

Risk factor phase

While under Broussais' conceptualisation the pathological link is direct and causal, under the risk factor paradigm it is probabilistic and its etiology only partially understood. Pre-pathological states are symptomless and exhibit mild levels of risk factors. In pathological states, quantitatively measurable risk factors are always probabilistically pathological. The logic of health seems to morally implicate a reduction of risk factors.



Figure 2.4: Relation of health and illness in risk factor phase

The notion of the 'normal' and the 'pathological' has seen dramatic shifts over the past two centuries of medical knowledge production. What is conceived of as normality has increasingly been encroached by a growing body of evidence linking quantitatively measurable physiological parameters to pathological outcomes. Such evidence was established on the conditions of probabilistic eventualities distributed across populations and over time. The polysemic nature of language also means that the semantics of normality are politically determined and distributed. Normality within the quantified-self movement,³⁵ for example, is something to be optimised and is somewhat uncomfortably reminiscent of earlier 'movements,' such as eugenics. Normality dictated by governmental regimes, such as public health and surveillance medicine, is more akin to something that needs to be retained and prevented from deteriorating. Health has been infused with normative status, produced through the promotion of health and self-surveillance as something that must be actively maintained by the preventive management and possible reduction of risk factors. Nevertheless, the introduction of risk factors rendered the maintenance of normality highly problematic. One could be symptomless and healthy and at the same time constantly at risk. Risk-factor epidemiology established the uncomfortable and probabilistically fragile, yet statistically significant, continuous link between health and pathology, over time and across statistical populations.

³⁵ People using a variety of sensor technologies to capture data about aspects of their daily life, such as sleep, food consumption, blood oxygen, typically with the aim of improving their 'performance' in relation to the average in the reference group (<http://quantifiedself.com>; ac. 18. 03. 2012).

Canguilhem's perspective on pathology and normality

An alternative conception of normality has been proposed by Canguilhem (1994). As previously discussed, risk factors are, however fragile, statistically pre-pathological. Canguilhem takes a broader perspective on understanding the supposed issue of deviations from the 'normal' and resorts to evolutionary theory based on "mutationist explanations" (ibid., 352). According to this dynamic view, biological normality is an outcome of the interactions between an entity and its environment over time. The normal is achieved when the entity exhibits an attempt to maintain, multiply and diversify itself and develops the capacities to adapt to environmental changes and, thus, achieve viability. In turn, the pathological is determined as inability to tolerate change (ibid.). In what follows, four different states of pathology are reconsidered from this mutationist adaptation perspective.

Symptomless risk factors state

Normality plus ability to creatively adapt to changes in environment = authentic normal as normative (e.g. non-aggressive prostate cancer).

Risk factors with mild symptoms state

Normality with signs of potential inability to adapt to changes in environment.

Pathological state

Normality minus the ability to creatively adapt to changes in environment = pathological normal (e.g. aggressive prostate cancer).

Unmapped space of risk factors

Not yet statistically or biologically evidenced as potential or actual risk factor.

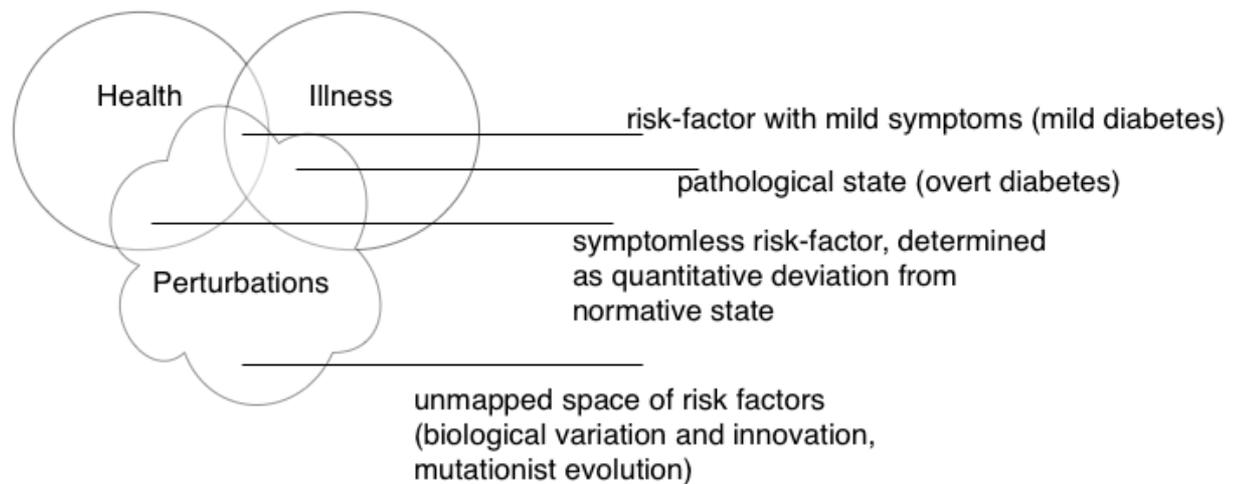


Figure 2.5: Canguilhem's perspective on health and normality

Such a view is liberating in the sense that it allows us to conceptualise symptomless risk factors as not necessarily and inherently linked to pathology, but also as long-term variation. Reflecting on the different levels of BMI and blood pressure considered normal among different national populations, such an adaptive mutationist evolutionary perspective allows us to understand physiological deviations beyond a normatively defined pathological state in need of preventive medical attention and intervention.

As a consequence of scientific advancement, some aspects of Canguilhem's concept might be due for a little revision. As indicated by recent genetic research (Kelley and Rinn 2012, 13), it is not so much the "structures and behaviours" of the human species itself that generate and in some cases succeed in mutational adaptation in evolutionary terms, but the environment itself. This seems to come in the form of retroviral molecular processes which act through the human and then, in some cases, become part of it. Integrating such molecular species-environment interactions allows us to retain Canguilhem's notion of the normal and the pathological; it also requires accepting the implication that, in some cases of deviations from the normal, identifying pathology can only be determined retrospectively over longer time-ranges, lending, thus, the preventive risk-factor-reduction logic an air of a somewhat myopic and human control-oriented strategy for long-term species survival.

It has been my intention to lay out and problematise the shifting notions of the 'normal' and the 'pathological' as they also constitute critical elements in assessing and making sense of medical knowledge. Whether and how these notions will yet shift again in the near future when medical research is

augmented by the developments of genetic-screening and pharmacogenomics as well as other data-driven approaches remains to be seen. The former two will not be addressed further within the scope of this work. In the following section, however, I will reflect and speculate on the potential role of the data-driven approaches in relation to conceptions of the normal and the pathological.

Data-driven medicine: Big Data

Big Data is a phenomenon operating in the realms of large scientific research projects in physics and astronomy³⁶ as well as big corporations. The first academic reference to the term 'Big Data' is by Weiss and Indurkha (1998). In their book on predictive data-mining they note that

very large collections of data [...] are now being compiled into centralised data warehouses, allowing analysts to make use of powerful methods to examine data more comprehensively. In theory, 'Big Data' can lead to much stronger conclusions for data-mining applications, but in practice many difficulties arise.

(ibid., xi)

Their book is mostly an applied discussion of methods for statistical evaluation of Big Data and preparing and handling Big Data volumes for such evaluation accordingly. Unfortunately, it does not refer or discuss the phenomenon of Big Data in any substantial way that would be relevant for our purposes here.

In his statistics and econometrics paper, Diebold (2012) similarly refers to an explosive data growth early on. By that he means the "explosion in the quantity (and sometimes, quality) of available and potentially relevant data, largely the result of recent and unprecedented advancements in data-recording and storage technology" (ibid., 13). In his discussion of two papers on novel economic Dynamic Factor Models being based on Big Data, he identifies that these novel models also operate with much larger numbers of indicators for regression analysis. Nevertheless, his discussion remains a technical one dedicated to the operative intricacies of his field. Following Diebold's argument, one would have to think that the development of Big Data

³⁶ See for example: <http://home.web.cern.ch/about/computing> and <http://openlab.web.cern.ch/publications/presentations/cern-big-science-meets-big-data> (ac. 15. 02. 2013).

came about as a natural consequence of Moore's law,³⁷ thus completely ignoring and obscuring other agendas that facilitated this development in the specific ways we have it today.

Other authors have further defined the term by adding descriptors such as "Volume, Variety and Velocity" (Laney 2003, quoted in Diebold 2012, 4) which, while certainly describing the phenomenon more accurately, still does not help better understand why it occurs. A paper by a group of leading computer science researchers from the USA representing the National Science Foundation and the Computing Research Association in their collaboration as the Computing Community Consortium only refers to "[A]dvances in digital sensors, communications, computation, and storage" that miraculously seem to "have created huge collections of data, capturing information of value to business, science, government, and society" (Bryant et al. 2008, 1), without any detailed discussion of what this value should be and under what conditions it can be generated. Before describing in further detail what Big Data is and why it occurs, it is useful to now turn to an example of Big Data analytics. This will help illustrate and understand the scale, operational mechanics and, importantly, some of its anticipated usefulness for contemporary society.

Big Data applied

Making use of large data sets are companies such as Google and supermarkets. Extensive customer profiles are assembled, containing socio-demographic, geographic and behavioural data as well as social graph information (Holmes 2007; Stalder and Mayer 2009). This, in turn, allows one to identify patterns among individuals or groups, which can provide the basis for more fine-tuned commercial or other³⁸ targeting of customer segments, or to drive and decide on product and service innovation (Tang et al. 2010). Big Data appears too broad a term when one does not consider the specifics of application contexts. Some of the generally hoped-for promises and likely implications of Big Data's seemingly attractive characteristics in relation to

³⁷ The observation that the number of transistors in integrated circuits and processors doubles approximately every two years (Moore 1965).

³⁸ The US army applies predictive analytics to anticipate the outbreak and specific development of Guerrilla types of warfare, see project SCARE (Spatio-Cultural Abductive Reasoning Engine), <http://www.economist.com/node/21553006>; ac. 15. 04. 2012.

healthcare can best be demonstrated by considering the Google Flu Trends (GFT) case.

For GFT, Google computes billions of search queries based on an automated method to identify early signs of potential influenza-related search activity in a region. Based on their model, this is then used to make predictions of actual flu trends in a region or country. Such early detection, the authors (Ginsberg et al. 2009) argue, is crucial for public health officials in order to prevent pandemic spread by resorting to relevant medical measures in a timely fashion and “may enable public health officials to mount a more effective early response [and] in turn reduce the impact of both seasonal and pandemic influenza.” This appears to be a valuable contribution, particularly when it comes to saving lives.

While the authors acknowledge that “panic and concern among healthy individuals may cause a surge in the Influenza-Like-Illness³⁹ (ILI)-related query fraction and exaggerated estimates,” they nevertheless suggest that “notable increases in ILI-related search activity may indicate a need for public health inquiry to identify the pathogen or pathogens involved” (ibid.). However, upon closer inspection of the methods applied, a number of questions and concerns arise, both in theory and practice. In order to unpack the implications such Big Data analysis may have, it is useful to highlight the key elements of the GFT model on the basis of which its relevance and trustworthiness are purportedly established. This will also allow for situating some of the issues with Big Data analytics in a wider context.

As the Google team describes the GFT model in their published article (ibid.), they built an automated system to identify influenza-related search queries, assuming no prior knowledge of Influenza-Like-Illness (ILI). The goal is to build a more comprehensive model for influenza surveillance than prior attempts (ibid.). To do so, they analysed hundreds of billions of individual searches submitted between 2003 and 2008. The model essentially builds on a line-of-best-fit (regression analysis) between most popular search queries in a specific week and US region and the probability that a GP visit during the same time and in the same region is ILI-related. The latter data is based on the existing flu surveillance system managed by the US Centre for Disease Control (CDC), which typically reports these figures with a time lag of 1-2 weeks. The GFT model ‘rewards’ queries which exhibit similar regional variances like those of the CDC data set. Based on this high correlation and real-time-ness of search query data as compared to time lag for CDC-data, early-warning signal quality is assumed.

The underlying assumption seems to be that most people initiate a search

³⁹ Influenza-Like-Illness is a flu-like syndrome, indicating the possibility of influenza.

query when they experience symptoms and check whether they are ILI-symptoms. It seems equally easy to imagine a scenario where some people initiating ILI-related search queries may have seen an ILI-related advert (assumed to be highly seasonal as well) or media article and preventively wanted to update their ability to spot symptoms of the specific seasonal flu in other people at work, on the street or in the family. Such seemed to be the case in the Bird Flu pandemic (H1N1) when the GFT model was indeed unable to correlate with pandemic-ILI data and needed to be adjusted for unexpected effects through extensive media coverage (Butler 2013; Cook et al. 2010).

From our perspective there seems to be, at least in theory, a significant potential difference between the intentionality of somebody making the effort to visit a physician for ILI-related purposes and the various reasons why someone might initiate ILI-related search queries for themselves or others, not to speak of the increasing numbers of people in the USA who cannot afford a visit to the GP.⁴⁰ Thus, while GFT seems able to crystallise general correlations between ILI-relevant GP visits and searches on temporal and geographic measurable vectors, it remains difficult to explain spikes in such search volume and why they differ significantly from actual CDC-ILI data as well as laboratory-confirmed influenza (Ortiz et al. 2011). This may be seen as a major drawback of this analysis method, as it does not offer sensible ways to further explore and make sense of findings and differences.

As highlighted by Ortiz et al. (ibid.), ILI itself is only an umbrella term for a “nonspecific syndrome that is not necessarily caused by influenza virus infection, but used for decades as an indicator of the burden of outpatient influenza illness.” While some see a practical use for GFT real-time detection data in calling in more staff or opening an empty hospital wing to allow more patients (Dugas et al. 2012), it remains to be confirmed that such decisions and even those more wide-scale ones by public health officials and epidemiologists in potential cases of pandemics and epidemics are indeed made and accounted for using GFT data.⁴¹

GFT as a case of Big Data analytics is exemplary for a number of specific ways in which Big Data is thought to provide its magic. For one thing, the sheer volume of data - billions of individual search queries - seems to cast questions of representativeness less relevant (Boyd and Crawford 2011). The GFT authors acknowledge this only insofar as statistical methodology requires them to state that “the correlations we observe are only meaningful across large populations” (Ginsberg et al. 2008). While such early indications may

⁴⁰ See <http://www.healthcare.gov/news/factsheets/2010/07/health-care-providers.html>; ac. 15. 02. 2012.

⁴¹ Personal communication with J. Ortiz, 20.3.2012.

nevertheless be valuable for rapid-response epidemiology, unfortunately, the authors highlight, and in contrast to traditional surveillance methods, no demographic data can be obtained from search queries (ibid.). Furthermore, the strength of correlation in light of such data volumes renders data outliers statistically less significant and, thus, less visible in the GFT interface.⁴² Consequently, the Big Data analyst can analyse at the comfort of a 30,000 feet perspective, hence deriving meaning not from decoding individual queries in context, as, for example, with ethnographic thick descriptions, but based on the emergence of patterns out of millions of data points. Furthermore, the question of whether, despite its historic existence, ILI is in fact a useful metric to compare with as indicated by the research by Ortiz et al. (2011) above remains to be answered.

Analysing at that 'altitude' or distance is further justified by relating Big Data to the 'real' through correlating it with grounded events,⁴³ the actual ILI-rated GP visits in a given time and geographic area. While such analysis may turn out significant correlations, as is the case in GFT, it obscures the question of their actual meaningfulness and relevance (ibid.) while assuming an unproblematic universal applicability (Goel et al. 2010). GFT has so far been applied to 29 countries with varying degrees of success (Cook et al. 2010); thus, while being an interesting approach, it may best be taken with some care.

Lastly, the insights generated are presented - at least to the public - as an end rather than as a means. By this I mean that data is not presented in a way that would allow one to make further context-specific sense beyond the findings churned out by GFT.⁴⁴ From a perspective of search as exploratory, situated and, thus, open-ended, this seems unnecessarily narrow and vexing.

A brief history of Big Data and why it occurs

While the massive growth of storage and processing capacities in digital computing are important contributing elements which have laid and continue to expand the necessary technological infrastructure underlying Big Data, they remain insufficient in fully explaining the emergence of such phenomena as

⁴² Personal communication with Richard Lewis, Computing Department Goldsmiths (19.3.2012).

⁴³ As Miller and Slater (2001, 1) argue, "[T]he internet as a meaningful phenomenon only exists in particular places" which then serve as "the only firm basis for building up the bigger generalisations and abstractions [...]"

⁴⁴ See <http://www.google.org/flutrends/se/data.txt>; ac. 15. 04. 2012.

GFT, and even less helpful for shedding light on what may be more fundamental driving forces for the rise of Big Data. Identifying such forces will help us start developing a more comprehensive understanding and definition of the term. This, in turn, allows us to identify and discuss the ways in which Big Data knowing may also be said to differ from the other data-based knowledges mentioned earlier.

A useful starting point is the larger shift from an industrial society to what has been variously named 'post-industrial society' (Touraine 1974), 'information society' (Machlup 1962) or 'network society' (Castells 1996). Even though there are important distinctions between these grand social theories, they all broadly describe the phenomena whereby substantial parts of contemporary economic, social and cultural activity manifests in and through the production, distribution and consumption of information. This shift has been particularly complemented and intensified not only by the aforementioned advances in computing and digital technology but also by the emergence of networked information and communication technologies, such as the Internet.

In an era of increasingly globalised economic markets, it is economically valuable and opportune to continuously stimulate, optimise and channel probable economic behaviour (Holmes 2007). Collecting, assembling and analysing data from customer transactions, which can then be further exploited for predictive marketing purposes, is inviting in a fiercely competitive economic environment. This is especially true as traditional broadcast channels do not support marketing campaign-tracking mechanisms very well. This has been corrected in the meantime, as these media updated themselves into their digital contemporary equivalents. Television has become Digital TV, or technically known as Internet Protocol Television (IPTV). Newspapers have also moved online. Both have become amenable to various regimes which aim to optimise 'user' attention with commercial value and, thus, the distribution of content considered relevant for these purposes.⁴⁵

Of course, 'economic drivers' for Big Data should also be understood in a much broader sense to include other sorts of surveillance infrastructures, be they for the smooth operating of urban environments (Goode 2011), or in the USA increasingly for attracting voters (Holmes 2007). Regarding the design of urban environments, architecture critic Sze Tsung Leong (2001) refers to this as "control space" (quoted in Holmes 2007) while others, such as geographer and urban researcher Stephen Graham, as 'Splintering Urbanism' (2001). As Holmes (2007) persuasively summarises it, "[T]he environment is over-coded with an optimising algorithm, fed by data directly coming from you." This has

⁴⁵ Such as personalisation, behavioural advertisements, tracking of interest and subsequent display of advertisements on other websites.

not only led to certain companies building up vast data repositories⁴⁶ about individuals and offering their clientele to ‘pre-emptively market’ to their customers as they transition between different life-phases,⁴⁷ but also continues to economically incentivise and intensify the specific ways such constant optimisation manifests in our daily interactions.

Another example of such over-coding is the more recent practice of persuasion-profiling. This is a marketing approach concerned with the optimal display of product offers. ‘Optimal’ here is understood as selectively displaying cues to which a potential customer will respond most desirably. As the inventors describe, in the case of a book this could consist in highlighting that ‘friends’ from one’s social network have also liked or bought a product, displaying endorsements and appraisal by other well-known authors, or, in the case of price-sensitive clients and only for them, offering a discount.⁴⁸ Considering Big Data in such a light, the difference between identifying a potential criminal, buyer or political voter seems almost of secondary importance, or merely one of potentially many attributes of a comprehensive profile. In summary, then, commercial and economic drivers for the development of Big Data can be explained by an intensifying of identifiable patterns of human behaviour which can be statistically analysed and consequently segmented and targeted to optimise the probability of profitable economic transactions or otherwise desirable behaviour.

Alongside the aforementioned economic drivers there is a parallel development in scientific and academic research. Similarly, the core motivation stems from the kinds of questions which can be researched thanks to the availability of data but also the capacity to compute them. The National Science Foundation summarises this well in their Cyberinfrastructure Vision for 21st Century Discovery Report:

Once used by a handful of elite researchers in a few research communities on select problems, advanced computing has become essential to future progress across the frontier of science and

⁴⁶ Acxiom, a US company, divides the entire US population into 70 demographic clusters, according to “age, estimated household income, presence and age range of children, marital status, home ownership status, estimated net worth and population density (Acxion.com; ac. 11. 04. 2013).

⁴⁷ See case of pregnant teenage girl whose father was unaware of his daughter’s pregnancy and was complaining to Walmart for receiving ads for pregnant women (<http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all>; ac. 01. 10. 2012).

⁴⁸ See www.persuasion-profiling.com/ and <http://captology.stanford.edu/>; ac. 10. 02. 2013.

engineering [...] Today's scientists and engineers need access to new information technology capabilities, such as distributed wired and wireless observing network complexes, and sophisticated simulation tools that permit exploration of phenomena that can never be observed or replicated by experiment.

(NSF Cyberinfrastructure 2007)⁴⁹

What seems of particular importance is their mentioning of a broader access to such infrastructure by a larger research community. The interest in and potential of Big Data for the humanities and social sciences research community, Manovich (2011) argues, stems from the diffusion of the stark limitations of up until recently standard data-supported research methodologies. As he explains, there was, on the one hand, 'surface data' containing few data points about many people. This was typically used in fields that work with quantitative methods, such as sociology, political sciences, economics etc. On the other hand, there was 'deep data' with a lot of data points about few people. This was typically used in the humanities, such as literary and film studies, history but also in non-quantitative schools, such as psychology and anthropology (ibid.). Manovich goes on to argue that statistics operates in-between the two methodologies with its concept of sampling technique, allowing the expansion of "certain types of data about the few into the knowledge about the many" (ibid.).

Unfortunately, as Manovich rightly points out, sampling comes at a cost; while it may in some circumstances allow for inferences about larger populations, as we will discuss later in more detail this is always based on a number of assumptions which we may never fully know whether they apply to a given population or not. Also, the sample is never representative of the actual behaviour of any individual or, indeed, all individuals studied; in this sense it only creates an image of average behaviour. To that extent, a researcher may well fail to appreciate all the different types of actual behaviour. For Manovich, the emergence of 'social media' has made it possible to eliminate the need to choose between surface and deep data. For him, this moment is well summarised in the following statement by Latour (2007, 2):

The precise forces that mould our subjectivities and the precise characters that furnish our imaginations are all open to inquiries by the social sciences. It is as if the inner workings of private worlds have been pried open because their inputs and outputs have become thoroughly traceable.

⁴⁹ See <http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf> accessed 1/8/2014.

Whether one agrees with the selection of terms used above, such as ‘precise forces’ and ‘precise characters,’ is open for discussion and will be addressed further shortly. Moreover, it would be safe to include ubiquitous computing as well as the use and widespread distribution of sensing and computational technologies built into everyday objects in Manovich’s picture of social media as a constitutive element for generating this possibility.

Nevertheless, Manovich’s argument is a useful one and resonates with earlier arguments made by John Tukey, a Professor of Science and Statistics in the second half of the 20th century. Tukey (1980, 23) argued that classic understandings of the functioning of science and engineering that follow a straight line from “[research] question => design => collection => analysis => answer” do not recognise how research ideas and questions are generated. He argues that this mainly happens through quasi-theoretical insights and the exploration of past data.

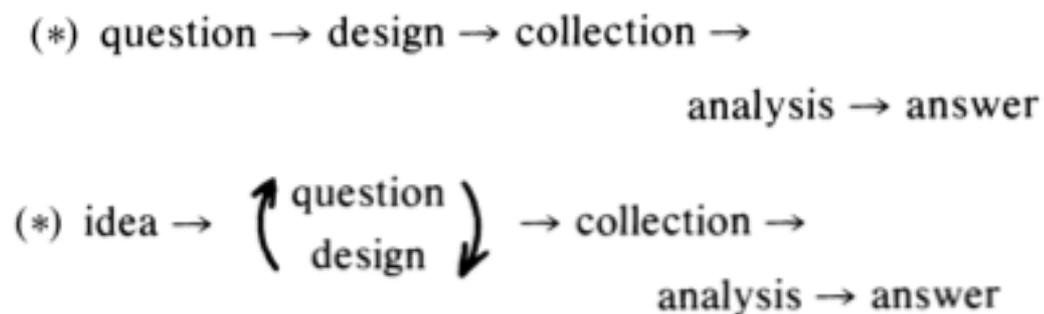


Figure 2.6: Tukey’s view on process of research idea and question generation

Most importantly, the ‘question’ and ‘design’ phases do not happen in a linear fashion but iteratively, feeding back into each other. For him, data exploration manifests and plays a crucial role in every step of the research process, from idea formation over research design to data collection, analysis and finding research answers. If we accept Tukey’s suggestion as still valid today, then extensive data sources such as Big Data may, indeed, signify a vast imaginable potential for asking many a research question. Taking a look at the contemporary research landscape, it is not difficult to find relevant examples of research organisations and projects that are uniquely focused on developing research based on Big Data, such as Manovich’s proposal for a Cultural Analytics (2008), Google’s Ngram Viewer⁵⁰ or a call for papers for a

⁵⁰ See <http://confluence-project.org>; ac. 10. 10. 2012.

2013 Digital Humanities Conference with the title 'Freedom to Explore.'⁵¹

In summary, the availability of Big Data is anticipated to relieve researchers from having to choose between and be satisfied with either surface or depth data. Besides physics and computing more generally, this is also attractive to both social sciences and humanities research in order to raise new research questions and develop new methodologies and practices. Such considerations are at the core of the recent discourse on Big Data, which aims to critically reflect on the social, cultural, economic and political implications thereof.⁵²

Reflections on Big data approaches to medical evidence

From the perspective of the medical knowledge paradigms discussed above, rather than marking a fundamental shift Big Data can be said to reinforce and extend the existing dominant statistical risk-factor paradigm. It does so mainly by augmenting and intensifying existing knowledge-seeking processes with large-scale computational capacities and broader sets of data considered. More specifically, it continues to complement direct, etiological understanding with the force of statistical significance, albeit now on a larger scale of data.

This can be demonstrated with the Google Flu Trends (GFT) case. For this comparative analysis, risk-factor research can be thought of as consisting of three broad phases: input, analysis and output. During the input phase, risk factors that are deemed relevant for the research are identified.⁵³ In the analysis phase, risk factors are methodically evaluated and compared to control groups with the aim of generating statistical evidence and probabilities as an output. It is important to notice here that, while the evidence is 'just' statistical and probabilistic, the underlying link between smoking and lung cancer is nevertheless thought to be causal in some cases evidenced in clinical tests.

The phases in the case of GFT are similar. What is different during the input phase is that the parameters considered are no longer directly linked in

⁵¹ See <http://dh2013.unl.edu/call-for-proposals/>; ac. 10. 10. 2012.

⁵² See <http://blogs.lse.ac.uk/impactofsocialsciences/2014/07/03/philosophy-of-data-science-series-evelyn-ruppert/>; ac. 16. 08. 2014.

⁵³ In the case of linking smoking to lung cancer, extensive interviews were conducted to understand potential causal links. To start with, the researcher suspected the use of tar on roads as a likely candidate (Doll and Hill 1951).

medical etiological terms but only used as measurable indicators of emerging pathological incidents. In the case of GFT, this is web search and click behaviour.⁵⁴ The analytic procedure applied in GFT also differs in one crucial way: in the absence of the necessity of a medico-etiological link, online behaviour is linked to emergent pathology by way of statistically proximate patterns compared to offline ILI-related pathological behaviour, that is people visiting GP's in relation to flue-like symptoms. When the patterns of assumed indicative online behaviour map neatly onto the relative offline behaviour related to ILI across a sample of the national landscape and within the time ranges observed, potential epidemic pathology is predicted.

Effectively, in the case of GFT a relatively broadly-intended, information-seeking activity assumes predictive capacity based on statistically-significant, homogenous past offline behaviour. In this sense, I argue that the risk-factor paradigm in the era of Big Data extends itself by further loosening the need for direct medico-etiological understanding in exchange for larger-scale statistical analysis of events which can be perceived as possibly indicative of anticipated pathological events. Given the absence of further exploratory capacity of the underlying data, its epistemological value is somewhat limited.

Such an approach to the production of risk-factor indicators brings with it a number of issues that need to be highlighted. It might be speculated that, just like statistical methods have evolved and increasingly acknowledged all sorts of potential and actual biases (Jadad and Enkin 2007, 29), a similar development might be anticipated for data science. Developing such an appreciation, though, hinges on accessibility and a relatively open conduct. Given the relatively undisclosed practice of Big Data analysis, it may take a while before a broader set of critical methodological concerns surface. By considering broader sets of potentially relevant data, Big Data projects may also identify correlations related to potential confounding effects of prior research studies that had to be satisfied with controlling a smaller set of variables. What, if anything, these correlations mean still requires a more fully extrapolated underlying theory so as to make sense of such data patterns as, for example, was the case with lung cancer. This resonates with Tukey's call for iterative feedback loops when exploring past data to derive quasi-theoretical insights for further exploration. The immediate question, then, becomes whether and who has the dedicated resources and required skills to weed through all those correlations (Kobielus 2012). Also, one may wonder what kinds of questions are being paid attention to and which ones may not seem worthwhile enough. While the current practice of the global

⁵⁴ In other cases it is weight-gain and overweight/obese people in one's social network that serve as probable indicators (Christakis and Fowler 2007).

pharmacological industry concerning orphan diseases may be one indication, there may also be speculated a potential for more fine-grained identification of long-tail effects (Anderson 2006).

From the genealogical discussion regarding the refinement of glucose-measuring tools from urine to blood-based for diabetes, further parallels can be identified with how the risk-factor paradigm has evolved and the implications this had on the conception and identification of pathology. An example is the case of measuring subtle irregularities in vital signs of prematurely-born infants up to 24 hours before outward symptoms become visible. It is asserted that, while such minuscule irregularities are not significant to trigger an alert, “physicians can provide critical treatment to infants up to 24 hours before the infection gets worse - making a life-saving difference” (Pittman 2012, 12; Blount et al. 2010).

As has been the case historically, the refinement of measurements and methods also leads to a reconsideration of the borders between health, pre- and mildly pathological states and pathology itself. From the material accessible,⁵⁵ it is unclear how exactly a normal state is defined. The authors only refer to future refinement by “reviewing the correlation of computed features with infants who actually develop nosocomial infection” (Blount et al. 2010, 117). From an ethical perspective, it is obvious that quantitative deviations from such a normal state could hardly be established via a randomised controlled trial. Following such a procedure would mean that the control group would not be given treatment, which in some cases would potentially lead to death in order to establish the causal pathological link. It, thus, has to be speculated that normality is derived as the average from the overall population of infants or premature infants itself. Pathology, in turn, seems to be defined as deviations from such averages. Unclear, however, is where and how the border between health and pathology is set. Thus, the number of treated false-positive premature infants seems to remain unknowable for the moment. It also remains unclear what, if any, iatrogenic effects such pre-pathological intervention may occasion and how this is accounted for.

Consequently, Big Data may be expected to identify endless risk-factor correlations, in some cases without offering the possibility to explore and interrogate them as seems to be the case with premature infants and GFT (Butler 2013). The latter illustrates that, while deviations from the ‘normal’ can be measured, they need not be signs of the pathological but in some cases simply false-positives that stem from implications of multifaceted interactions

⁵⁵ The authors refer to a future publication regarding details of the nosocomial intervention of Big Data.

and other media influences. As Google can also capture such media influences via their Google News service, it remains to be seen how they will avoid to overly enthusiastically self-censor such potential media effects. As of this writing, there do not appear to be strong reasons to believe in a relatively open conduct for Big Data, given that APIs to this vast data pool are highly filtered (Boyd and Crawford 2011). This is hardly surprising as much of the immense stock-market capitalisation of corporations may be said to rely on the high hopes that these corporates will exploit the data for their own benefit and on their own terms. Lastly, given the availability of evidence, the step towards intensified policing of certain behaviours seems a small one for some.⁵⁶ While not within the scope of this research, the role of Big Data in surveillance practices more broadly has recently been identified and problematised.⁵⁷

Summary

As demonstrated in my discussion so far, the statistical techniques for establishing probabilities emerged historically as part of a regime for reform and medical policing. They have established a link between quantitative measurement of physiological - and increasingly behavioural - parameters and probable pathologies over time and across populations. However, it may be said that they have never been primarily intended as a source of information for meaningful decision-making for and by an individual. To that extent, they aim to speak of populations, evidently with little concern for specific distinguishing characteristics among individuals, apart from demographic aspects. It seems easy to see an improvement of 2% in the probability for men to detect the likelihood of prostate cancer with a PSA-test as a truly meaningful figure from the perspective of a public health manager. However, from the perspective of an individual, the improvement of probability figures from 42% to 44% hardly provides a substantially more meaningful basis for deciding on taking the risk of lifelong side-effects from a biopsy. In the worst case, only to find out that one has figured as a false-positive finding in the earlier PSA-test statistic. Thus from the perspective of a potentially affected individual rather than that of a public health manager, probabilities seem only partially useful for

⁵⁶ Doctors demand to withhold treatment from obese patients who fail to lose weight as data correlations reveal that the treatment seem less effective in those cases and, thus, tax-money is being spent less effectively. See <http://bit.ly/ICc7re>; ac. 15. 02. 2013.

⁵⁷ See <http://bds.sagepub.com/content/1/2/2053951714541861.full>; ac. 18. 08. 2014.

decision-making. As Rose argues:

Normativity now becomes a matter of normality, of social and moral judgements about whether particular lives are worth living [...] The judgements of probabilities and of risks that have become central both to experimental and clinical practice inescapably connect to the judgements of value that are placed upon different forms of existence and the logics of treatment they mandate. What is normality at the level of the genetic code? Is what is optimal for the population necessarily optimal for the individual?

(Rose 1998, 165)

The paradigm of risk factors and, thus, the meaningfulness of probable risks is epistemologically characterised by a sharp contrast between a population view and an individual perspective. Meaning from a population view implies putting a higher value on experimental practice for improving the average health of the population with regard to a specific aspect. This takes place at the cost of framing healthy people as at risk of pathology, which comes as a necessary consequence of the inherent methodological limitations. Despite the questionable meaningfulness of risk factors at the individual level, such statistical modulations become active mediators (Hacking 2006). The specific ways in which probabilities are useful and problematic to individual health decision-making will be analysed and discussed in detail in the following chapter.

Chapter 3

Shared medical decision-making and decision support

Introduction

This chapter builds upon the biopolitical concerns of evidence-based medicine (EBM) I have analysed in the previous chapter, such as the changing possibilities and responsibilities projected onto people in sustaining and promoting their health. I will do so by considering the assumptions underlying the design of distributive mechanisms for risk information concerning cancer.

One way such self-responsibility is claimed to be supported is by means of decision aids (DA) as a tool in healthcare procedures involving shared decision-making (SDM). Decision aids come in the form of videos or pamphlets and are developed to support people in making an informed and unbiased decision in a medical situation. This is thought possible by providing a balanced view of benefits and harms of treatment options according to the current evidence-base. Decision aids, thus, consolidate and operate as the distributive mechanism by which evidence-based medicine is rolled out to the population. In turn, they constitute an object of critical concern for this research, both in terms of theoretical analysis as well as design practice. Based on the analysis of the limitations underlying the current design of DA's, the subsequent practice-based design project aims at redesigning such decision support.

The chapter will first define SDM, consider its genealogy and discuss the principles based on which decision aids are designed and evaluated. I will do so first by analysing two systematic reviews of the performance of DA's and the criteria based on which the evaluation was conducted. Furthermore, I will analyse the specific ways in which risk information is presented and the implications this has on individuals' sense- and decision-making. This renders more transparent the underlying assumptions and design decisions materialising in the design of contemporary DA's. It is found that, while presenting risk information with the use of natural frequencies⁵⁸ may facilitate its immediate readability, I argue that such an approach simultaneously further obscures the fundamental uncertainties underlying the generation and

⁵⁸ This is a term introduced by Gigerenzer and Hoffrage (1995). They argue that human cognitive processes have evolved along with the sequential counting of events rather than with probabilities. Thus, according to evidence collated by them, humans (as well as animals) perform much better when information is presented this way.

applicability of such information, in particular at the level of the decision-making individual.

In the second part of the chapter I analyse the ways in which such informational uncertainties relate to patient autonomy and, in turn, to SDM processes. Specifically, I will suggest reconceptualising patient autonomy beyond punctate decisions.⁵⁹ This will also include the larger healthcare context, socially-derived standards and norms that influence what is considered responsible individual behaviour and the individualisation of risk more broadly. Subsequently, I argue that the processes for constructing preferences and personal values are of particular interest for this research. Unfortunately, relatively little prior research exists as research into SDM thus far has not fully addressed this aspect.⁶⁰ The chapter closes with suggesting reconsidering patients as active inquirers and patient autonomy more as a collaborative knowledge practice, in particular with regard to the construction of preferences and values.

Shared Decision-Making

According to the Informed Medical Decisions Foundation, one of the earliest organisations⁶¹ to research and develop tools for shared decision-making in the USA,

Shared decision-making is a collaborative process that allows patients and their providers to make healthcare decisions together, taking into account the best scientific evidence available, as well as the patient's values and preferences.⁶²

⁵⁹ Kukla (2005, 35) refers to punctate decisions as seen by ethical theorists as those discrete choices independent of any implications in a person's larger health care situation.

⁶⁰ This is argued by Epstein and Peters (2009, 196) as well as inferred from the Cochrane systematic reviews and the International Patient Decision Aid Standard evaluation criteria underlying them as discussed later.

⁶¹ Established in 1989 by John Wennberg and Albert Mulley. Wennberg, as encountered in chapter two, is the epidemiologist who initially identified the unwarranted small-area variation in medical practice (1982). Mulley will be discussed with regard to the patient preference-construction process in this chapter. The goal of the organisation is to advance evidence-based, shared decision-making by supporting research.

⁶² Informed Medical Decisions Foundation, <http://informedmedicaldecisions.org/what-is-shared-decision-making/>; ac. 01.09.2013.

Shared decision-making evolved partly from critical voices and social movements “challenging the paternalistic power that doctors exercised over their patients and their lives” (Rose 2006, 10) as well as from the findings on geographic variation of medical services by Wennberg (1973, 1982) that led to the promotion of evidence-based medicine discussed in the preceding chapter. The latter are the geographical variations in clinical practice for what was seen as essentially identical medical problems (Chassin et. al 1986).

An example of such geographic variations is that in some regions of the USA essentially no women underwent lumpectomy surgery⁶³ for early stage breast cancer treatment, whereas in others the figure was 50%.⁶⁴ While the variation could theoretically also stem from differences in patient choices, the statistical variances did not seem to support this credibly as the patterns seemed geographically concentrated, whereas variation due to patient choice, it was thought, would have led to more randomly distributed spatial patterns (Dartmouth Atlas Project Topic Brief 2007). The variation of clinical practice, thus, was found to be attributable, among other things, to local medical opinion, also dubbed ‘practice style factor’ by Wennberg (1984). By this, Wennberg hypothesised that the geographic variation in medical services used stems from differences in clinical judgment by physicians concerning favourable treatment forms. While some variation seemed acceptable due to the complexity of medical situations, the stark overall variations were also of concern with regard to rising healthcare costs (Epstein 1990).

The findings led to the formation of what has been named the ‘Outcomes-movement’⁶⁵ (ibid.). The motivation and goals of the movement included cost-containment as well as a consideration of a broader range of outcomes relevant for assessing treatment outcomes and health (ibid.). The latter has had a direct influence on the emergence of SDM. Research was conducted to start assessing different treatments and their outcomes. In one such research project concerning benign prostatic hyperplasia,⁶⁶ the researchers developed a model based on expected utilities to compare a watchful waiting approach to surgery intervention (Barry et al. 1988). Watchful waiting, as the name indicates, requires regular inspections of symptom

⁶³ Surgery to remove lumps from the breast.

⁶⁴ Similarly stark discrepancies were identified in other medical treatment contexts (Chassin et al. 1986).

⁶⁵ With this, Epstein (1990, 266) refers to a variety of organisations, such as the Health Care Financing Administration or the newly established Agency for Health Care policy and Research, which directed their activities or launched programmes for assessing “the effectiveness of medical interventions and developing guidelines for medical practice.”

⁶⁶ An enlargement of the prostate.

progress before surgery is performed, whereas surgical intervention is an act of preventive surgery with a probability of implicating various side-effects, such as incontinence and difficulties with erection (BMJ Group). ‘Utilities’ were here determined as measurable changes in quality of life months⁶⁷ as a result of these treatments.⁶⁸ Based on this quantitative comparative model, the research concluded that there is a great degree of variance in expected quality of life utilities following the different treatments and, thus, “patient preferences should be the dominant factor in the decision whether to recommend prostatectomy” (Barry et al. 1988, 3010). These findings were the launching pad for shared decision-making procedures and started to establish the relevant conceptual basis for SDM, including the distinction between ‘preference-sensitive’ and ‘effective’ care (Dartmouth Atlas Project Brief 2007).

Preference-sensitive care is a healthcare situation with several treatment options which are considered to have equivalent outcomes. Effective care situations, in contrast, have a single best-evidenced medical practice (Wennberg 2002). Equivalence of healthcare outcomes is defined and measured in terms of indices, such as life-expectancy and quality of life months as discussed above. SDM, thus, is the procedure aimed at supporting patients in making such preference-sensitive medical decisions based on their values and preferences.⁶⁹

The principles it follows and goals it aims to achieve thereby are threefold: firstly, ensuring clinical practice is fully informed by best available knowledge based on EBM; secondly, in preference-sensitive care situations, involving and eliciting patient preferences and values; lastly, reducing variation in clinical practice resulting from what was seen as “irrational and uncertain application of medical knowledge and experience” which, in some cases, was argued to have led to detrimental effects for patients (Chassin et al. 1988). Thus, SDM is also fundamentally different from the principle of informed

⁶⁷ Health-related Quality of Life (QoL) is determined by way of surveys among healthy and ill people, identifying five categorical aspects which have the most significant impact on QoL including mobility, self-care, usual activities (work, study, housework and family or leisure activities), pain/discomfort and anxiety/depression. For example, having “some problems walking about; some problems washing or dressing self; some problems with performing usual activities; moderate pain or discomfort, or being moderately anxious or depressed” is calculated to reduce the quality of life by about 50% (see <http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-3l-value-sets.html> ac. 01.03.2014).

⁶⁸ The data for such utilities were sourced from the medical literature, patient interview studies and Medicare claims data (Barry et al. 1988).

⁶⁹ See also Informed Medical Decisions Foundation, <http://informedmedicaldecisions.org/what-is-shared-decision-making/>. ac. 12.5.2014

consent. The key difference between the two is that, in the case of the latter, the physician evaluates available treatment options for a patient, selects one and is by law bound to get a patient's consent to it. In contrast, with SDM the patient makes an informed choice based on the options explained by the physician, including an understanding of probable risks and benefits (Center for Public Policy and the Social Sciences 2011).

Among the tools that are used to support SDM procedures are decision aids. Decision aids are used to frame the decision to be taken by explaining the available treatment or therapy options in balanced and unbiased ways. Furthermore, the decision aid highlights existing and absent knowledge about probable risks and benefits implicated in the various treatment options. Lastly, some decision aids may include a discussion of what the likely experience and possible side-effects of the treatment options available may be.⁷⁰ Decision aids have been developed for a number of medical situations, ranging from whether to take a prostate specific antigen (PSA)-test⁷¹ when concerned about prostate cancer, to the contextualization of end-of-life care, such as whether one should stop treatment that prolongs life.⁷² Such decision aids, then, highlight and provide information about the condition itself, different treatment options available, probable implications on life-expectancy, effects on quality of life as well as strengths and shortcomings of available evidence. Decision aids come as paper-pamphlets, videos or online tutorials. An example of a decision aid concerning the question of whether or not to take a PSA test is attached in Annex 1 and will be analysed in the section 'Analysis of a decision aid.'

Status quo of decision aids

There have been various assessments concerning clinical decision support both for physicians and patients. The following section will consider the evidence of how patient decision aids perform and the issues identified which impede the efficacy of this approach and set of techniques. Furthermore, I will review and discuss the assessment criteria used in the evaluation of the performance and effectiveness of contemporary decision aids. This will help to further characterise and define decision aids through the expectations on which the assessment criteria are based.

Two systematic reviews for the evaluation of the performance of decision

⁷⁰ See also Informed Medical Decisions Foundation, <http://informedmedicaldecisions.org/shared/decision-making-in-practice/decision-aids/> accessed 12.5.2014.

⁷¹ See <http://sdm.rightcare.nhs.uk/pda/psa-testing/introduction/>. ac. 03.08.2014.

⁷² See <http://decisionaid.ohri.ca/AZlist.html>. ac. 03.08.2014.

aids were performed by the Cochrane Collaboration in 2003 and 2009 respectively. While the findings reflected here are based on the 2009 review, they are largely congruent with the earlier one. Significant and relevant deviations will be highlighted and discussed. One such significant difference between the two reviews is that the 2009 review was also able to consider the use of probabilities in decision aids. In the 2009 review, a total of 55 randomised controlled trials were considered which evaluated the “efficacy of decision aids for people facing difficult treatment or screening decisions” (Cochrane Collaboration 2009, 1). The review found that, when compared to common care interventions, decision aids performed favourably by improving people’s knowledge about options available, reducing decisional conflict with regard to feeling uninformed and unclear about personal values, and lowering the percentage of people undecided about which option to choose. Furthermore, the inclusion of probabilistic information about risk also led to more accurate risk perceptions, particularly when presented in quantitative form (ibid., 43).⁷³ The review also found that decision aids continue to dampen “enthusiasm for major elective surgery [...] in favour of more conservative options”, as well as accounting for a decreased preference for PSA testing (ibid., 43).

The review results, however, also point out that decision aids do not perform better than conventional care practices (that is consultations without the support of decision aids) when it comes to satisfaction with decision-making, anxiety and health outcomes such as quality of life (ibid.). The assessment of the decision aid’s performance was based on primary outcomes criteria as developed by the International Patient Decision Aid Standards (IPDAS) Collaboration⁷⁴ as well as secondary outcomes concerning aspects of behaviour (decisions), health outcomes (quality of life, anxiety) and the healthcare system more generally (e.g. patients and physicians’ satisfaction, costs, litigation rates) (Cochrane Collaboration 2009, 4).

Among the primary outcomes criteria figures evidence “that the decision

⁷³ The underlying RCT’s have evaluated the performance of decision aids by using a questionnaire to assess patient knowledge. The questionnaire consisted of 36 statements to be answered by ‘true’/ ‘false’/ ‘unsure’ responses. Furthermore, to assess risk perception based on quantitative probabilities, 4 items with 4 multiple-choice options were included (Whelan et al. 2004, 437). Importantly, the term ‘risk’ here is understood in the narrow sense of the distribution of known potential outcomes following a treatment, such as incontinence. This is different from uncertainty, which relates to unknown implications as well as their likelihood of occurrence (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1553508/> ac. 23.07.2014).

⁷⁴ A collaboration of more than 100 researchers, practitioners, patients and policy-makers from 14 countries (Cochrane Collaboration 2009, 3).

aid improves the match between the chosen [treatment/ screening] option and the features that matter most to the informed patient” (ibid.). Unfortunately, there were no trials evaluating the criteria of whether decision aids help patients understand that values affect the decision, or discuss values with the practitioner” (ibid., 2). This latter aspect, which will be discussed in the section ‘patient preferences and autonomy’ in detail, seems important in the context of preference-sensitive care, such as cancer screening.

The findings of the review that decision aids are capable of reducing decisional conflict and developing more accurate risk perceptions when presented in quantitative form seem favourable outcomes. However, and surprisingly, despite these findings, decision aids seem unable to improve satisfaction with the decision-making process or relieve anxiety. In order to better assess what to make of these findings, it is necessary to more deeply inspect the ways in which the presentation of risk information is addressed in decision aids, and how this might be connected to satisfaction with the decision-making process and anxiety.

One of the studies included in the Cochrane Collaboration systematic review is that of Whelan et al. (2000) in which they tested a decision aid that focused on the interaction and communication between physician and patient concerning breast cancer treatment options. From reviewing their decision aid that can be viewed online⁷⁵ it is noticeable that quantitative probabilities are presented in the form of natural frequencies such as ‘1 out of 10’ or, at another place in the DA, as ‘5 to 10 out of 100’ with regard to the possibility that not all of the cancer might be removed and subsequent surgery may be necessary. This points to the question of the form in which probabilities are presented in DA’s, whether as percentage probabilities (e.g.10%), natural frequencies (e.g. ‘10 out of 100’), verbally or visually represented, and the effects this might have on risk perception.

Research on this question stems from psychology. Looking at two recent prostate cancer decision aids (BMJ Group; Health Dialog) both make use of verbal descriptions as well as natural frequencies, but no use of visual representations. Textual descriptions are complemented with natural frequency statements for precision. This seems to follow previous research findings which indicate that textual descriptions of probabilities alone may cause different interpretations (Wallsten et al. 1986). By using natural frequencies instead, a number of potential problems and biases may be prevented. Among them are framing effects and ratio bias or denominator neglect. Framing effects refer to the practice of highlighting the positive or negative effects (whether probable or actual) only instead of giving a balanced

⁷⁵ See <http://jco.ascopubs.org/content/17/6/1727/F1.large.jpg>. ac. 10.07.2014.

view of benefits and risks, such as for example stating a 97% survival rate. Ratio bias and denominator neglect refer to the perceived difference in risk between stating risk as ‘1 out of 10,’ ‘10 out of 100,’ and ‘100 out of 1,000.’ Typically the larger the numerator, the larger the perceived risk (Rothman and Kiviniemi 1999). There is little research on assessing the effects of visualising uncertainty. From reviewing a variety of material (Lipkus 2007; Lipkus and Hollands 1999; Spiegelhalter et al. 2011) it rapidly becomes obvious that, with regard to the visualisation of risk probabilities, these are often presented in a clustered style (see below).

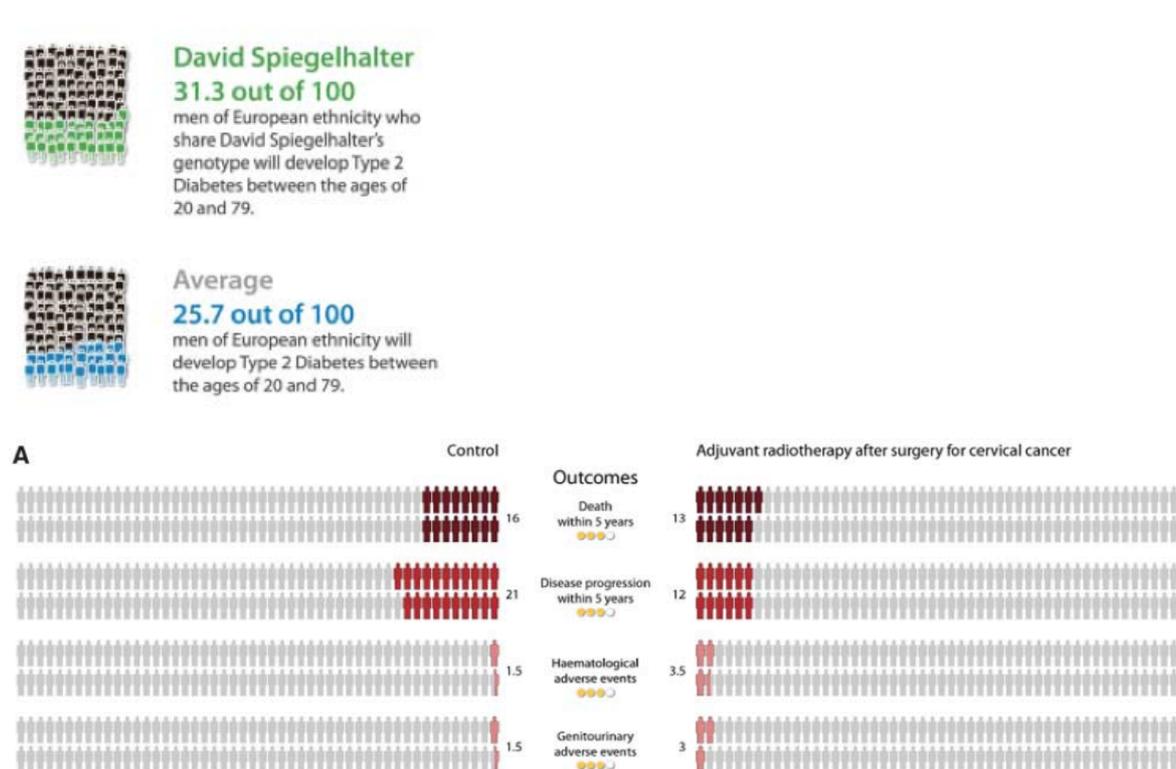


Figure 3.1: Visualising uncertainty about the future (Spiegelhalter et al. 2011)

One study from the discussed review that is of specific interest for this research is that by Ancker et al. (2006), in which the typically clustered style of visualising risk probabilities has been compared to a scattered style. What the study found was that, while a scattered visualisation style renders apparent the inherent unpredictability of such risk factors as they relate to individuals, particularly people with a low degree of numeracy found it difficult to assess and compare magnitudes of risk. Typically, risk visualisations privilege the display of magnitudes of risk at the cost of the unpredictability of risk factors.

From the perspective of an individual person, privileging magnitude of risk as opposed to unpredictability seems partly a questionable design

decision because communicating and understanding unpredictability is at least as important to prepare for a variety of possible health outcomes, irrespective of the magnitudes of risk factors which at the individual level have “limited applicability” (Politi et al. 2007). In turn, the question of how to design for an understanding of unpredictability and uncertainty is an interesting one and will be discussed further in the sections below as well as in the following chapter.

From their review of evidence on styles and effects of uncertainty visualisations, Spiegelhalter et al. (2001, 1399) advise that healthcare advisors “use multiple formats because no single representation suits all members of an audience” and “perhaps the greatest challenge is to make a visualisation that is attractive and informative, and yet conveys its own contingency and limitations.” As research by Gigerenzer et al. (2007) demonstrates, the capacity to understand statistically-derived risk information, and in particular the degree of uncertainty pertaining to it, is rather low amongst both physicians and patients. The inability to understand and clearly communicate health statistics renders informed shared decision-making an impossibility. Among other aspects, their research indicates that both physicians and patients have problems understanding conditional probabilities, because such statements are linked to different reference classes. Instead, their research findings suggest the use of natural frequencies, which refer to base rates,⁷⁶ and, thus, aim at preventing people from committing errors due to the base-rate fallacy.⁷⁷

To give an example in the context of breast cancer screening, a conditional probability may be formulated as follows: “The probability that a woman has breast cancer is 1% (prevalence). If a woman does not have breast cancer, the probability that she nevertheless tests positive is 9% (false-positive rate).” Using natural frequencies instead, the above would be formulated as follows: “Ten out of every 1,000 women have breast cancer. Of the 990 women without cancer, about 89 nevertheless test positive.” While natural frequencies seem a means to communicate risk information that is less prone to misunderstanding, it may be argued that simultaneously such an approach neglects addressing the more fundamental uncertainties underlying such risk information.

Epidemiological risk information for preventive medicine maps a distribution of possible outcomes according to their probabilities. However, as discussed in chapter two, such mapping is predicated on a set of assumptions which render the positive predictive value of risk information at the individual level largely uncertain. These assumptions include, among others, whether the

⁷⁶ A base rate in statistics is the categorical population to which an enumerator refers to.

⁷⁷ This is the error in thinking when ignoring the base rate to which an incidence rate refers. For example, the effectivity of a treatment of 100 in 1,000 versus 100 in 100,000 is vastly different.

individual decision-maker is well represented within the sample population based on which the risk factor was derived, whether all of the deterministic factors that contributed to pathology in the reference population will remain stable and, thus, apply to future populations, uncertainty about the lead time to pathology and its course, as well as the notoriously low discriminatory ability of epidemiological risk factor at the individual level. Uncertainty, thus, arises as a consequence of the difficulty to assess whether a person is not just statistically but actually at risk, as well as the distribution of probable outcomes. Apart from this technical view of uncertainty, there are other notions of uncertainty which either concern the physician-patient interactions (Politi et al. 2007, 690)⁷⁸ or relate to patient values and preferences, that is to say the second half of the definition of shared decision-making above, to which I turn now.

Patient preferences and autonomy

Patient autonomy is an important element of shared decision-making. This is because, in contrast to informed consent, in shared decision-making concerning preference-sensitive care situations the patient herself is expected to demonstrate the necessary level of comprehension so as to be able to choose and decide on the best course of action. From an ethical perspective, it is suggested that such autonomy is defined as “the ability to make and act upon free, informed decisions resulting from capable and uninfluenced deliberation” and “the relevant unit of autonomy are punctate decisions“ which refer to distinct choices unrelated to a patient’s larger health care situation and context (Kukla 2005, 35).

As research on the effects of communicating uncertainties indicates, the effects thereof can be beneficial as well as detrimental. The benefits concern the capacity to make such informed decisions. On the detrimental side, providing exceedingly complex information in combination with the suggested simplifying strategies broadly applied in handling risk information (Kahneman and Tversky 1974) can, in some situations, lead to avoiding decision-making altogether (Politi et al. 2007, 688). As Kukla (2007) and Epstein and Peters (2009) point out, in some medical situations where the problem, options and consequences are clear, eliciting patient values and preferences is relatively uncomplicated. However, in circumstances which exhibit “high-stakes, and uncertain situations with potential outcomes that have not been considered or cannot be imagined” (ibid., 195), eliciting preferences and values is much

⁷⁸ Physicians and medical trainees in some cases fear that they are perceived as inadequate or ineffective, which affects their willingness to disclose uncertainty (Politi et al. 2007, 690).

more difficult. In such situations, they argue, preferences are constructed rather than simply elicited.

While there is little research on the effects of communicating uncertainty instead of just risk (Politi et al. 2007), research does indicate that “patient satisfaction is affected by the manner in which the physician handles uncertainty, not whether or not he or she presents uncertainty” (ibid., 690). In turn, patient satisfaction and trust can be high when a physician is confident with uncertainty and it is shared in a collaborative way between physician and patient. Some suggest that medical practitioners focus on helping patients cope with uncertainty rather than just understanding it (ibid.). This points to questions regarding the conceptual and operative understanding of patient autonomy; it is worthwhile further inspecting this before considering the ways in which the construction of patient preferences can be meaningfully supported.

As Kukla (2005 and 2007) points out, understanding patient autonomy is best achieved when considering not just the punctate decisions, but also the larger health context of a person. With reference to prenatal healthcare, she illustrates well that ethical concerns have focused mostly on “moments of crisis or discrete choice [...] such as decisions to have or not have prenatal tests, abortions, or treatment interventions” (Kukla 2005, 36). However, as she argues, typical prenatal healthcare is seldom made up of such tests and decisions, and is, instead, best characterised as “routine, ongoing activities, some carried out by the woman herself under medical supervision, and some carried out in the clinic” (ibid.).

This includes all sorts of self-monitoring, following certain health, food and exercise regimes, as well as organising to undertake regular medical surveillance tests, and in case of symptomatic events even check themselves into emergency care (ibid.). In short, “pregnant women are expected to be active, self-disciplining participants in their own prenatal healthcare” while, importantly, such prenatal healthcare practices are governed by socially-derived standards in terms of “what we take them to be responsible for doing” (ibid.). Thus, Kukla maintains, in essence the sociocultural and medical norms governing prenatal healthcare are largely pre-existing, hence found rather than being self-determined by pregnant women. Importantly, and socially, pregnant women are not just expected to be merely compliant with such norms, but, instead, “take them on as their own” (ibid.). As Kukla rightly argues, some people would likely deem a pregnant woman abstaining from alcohol not due to her own sense of responsibility but only because her doctor said so somewhat suspicious. In summary, the example of prenatal healthcare and the social and medical norms by which it is guided is exemplary of larger ongoing

healthcare activities, which for Kukla “take the form not of crisis management and punctate decision-making, but of ongoing practices, including large amounts of self-management and surveillance [...]” (ibid.).

It seems relatively simple to see similar forces at work in cancer screening, including public campaigning by pro-screening organisations, such as Movember,⁷⁹ media presence of Hollywood celebrities suggesting to reduce one’s personal cancer risk by preventive treatment, and the way such approaches build on and further propel the notion of the individualisation of risk and responsibility. Underlying the dominant, contemporary epidemiological research is a focus on individuals and understanding their potential pathological future by way of genomic mechanisms. On what grounds it is rational and acceptable for an individual to decide “to lower his ‘individual risk’ of colon cancer from a high risk of 32 in 10,000 in 5 years to a low risk of 8 in 10,000 in 5 years” is open to discussion (Rockhill 2005, 126).

Not only is it unclear how patients respond to personalised risk estimates (Politi et al. 2007, 692) but also there is no understanding of the possible “consequences of imposing massive ‘risk awareness’ among healthy individuals through the general population” (Rockhill 2005, 126), although narrative problematisations (Sontag 1977) as well as indicative evidence (Milton 1973) exist. A side effect of such a focus on individualisation is that social and economic aspects that affect public health are generally “denigrated as ‘political’ or ‘social,’ rather than scientific concerns,” hence deemed irrelevant for the epidemiological agenda (Rockhill 2005, 127). This resonates strongly with earlier such concerns as discussed in chapter two (Rosen 1974, 67). Also, as Politi et al. (2007, 692) propose,

exactly how to use risk estimates and risk prediction tools to improve and inform individual treatment decisions, while acknowledging and communicating their limited power to predict individual futures, is a critical challenge that will become even more important as new disease biomarkers are discovered.

Following such considerations, the construction of preferences is deemed to be of particular importance within the specific context of prostate cancer screening. As discussed earlier, the construction of preferences can be influenced by various biases, such as the ways in which information is presented or framed. Within the medical context, the physician-patient encounter may be said to offer the potential to support such preference-construction processes. Unfortunately, little research exists beyond carefully

⁷⁹ See: www.movember.com. ac. 02.08.2014.

controlled experiments to better understand the dynamic interactions and relationship between physician and patient in this regard (Epstein and Peters 2009, 196). As Epstein and Peters further point out, current clinical guidelines and shared decision-making do not fully address questions on what role physicians should take in the construction process and how to best support the patient within this process to best cope with uncertainty. Nevertheless, Epstein and Peters suggest that respecting and responding to patient preferences involves shared deliberation that moves beyond mere information provision to helping patients explore, question and construct preferences. They suggest that, while the physician-patient relationship provides a framework for exploring preferences, other medical professionals and family members involved can be helpful as well as decision aids, which are “sensitive to default options, framing, and ordering effects [and] encourage deeper discussions with family members and clinicians” (ibid.).

Such an approach to deriving patient preferences may, in turn, best view the patient as an ‘active inquirer.’ As encountered in chapter one, the Internet provides a very rich source of medical information, which may also affect the physician-patient relationship. In contrast to ethical discourse, where, according to Kukla (2007, 28), patients are often framed as passive recipients of medical knowledge by medical professionals, patients may engage in active inquiry before they visit a medical professional (Zickuhr 2010).

Due to differences in health and statistical literacy such inquiry may be subject to less than perfect judgment regarding the quality of information found. Given the technical complexity and typical unfamiliarity of laypeople with the medical knowledge domain, one of the key tasks for a patient inquirer is discerning which sources of information to trust in order to develop an understanding of the issues concerned. Reliance on select experts as well as other concerned participants in the field can, thus, be considered as operatively reasonable within this specific context. Consequently, autonomy may be less understood as derived from ‘pure’ individual learning and knowing in every detail, but more as a kind of collaborative knowledge practice, especially concerning the construction of preferences and values in the context of a novel health situation. Within the context of lay medical inquiries, Kukla suggests to reposition autonomous inquiry as “skilled and competent coping in the course of our investigational practices - our attempts to discern, know, predict, and make appropriate guesses and decisions” (Kukla 2007, 30). For her, such coping and the capacity for autonomous inquiry practically materialises and depends in part upon “our social, institutional, and material position” (ibid.).

Such collaborative practices, however, can be subject to a variety of influences and constraints. For example, there is a history of discriminatory practices concerning patients that question or challenge medical authority (ibid., 31, 33). Furthermore, the contemporary healthcare system is highly specialised, making it difficult in some cases for a patient to get a consolidated view of his health situation (ibid.). Also, not all medical interventions require informed consent (Kukla 2005, 36) and may, thus, simply be characterised as 'routine procedure' and be less likely to invite critical inquiry. This is despite the fact that medical history has witnessed many a routine intervention which over time turned out to be of little effectiveness at best (ibid., 31).

According to a classic model of the physician-patient relationship (ibid., 32), physicians are responsible for providing the medical technical expertise or 'facts,' whereas the patient's role is to be an expert on their own values. Medical decisions are, then, based on applying the patient's values to the physician's 'facts.' This seems an overly simplistic model. The idea that medical facts are always fully objective is naive at best, as I have discussed by way of my analysis of the values and biopolitics inherent in the development of medical epistemology in the previous chapter. Furthermore, patients may indeed not yet be experts regarding their values when faced with a novel medical situation and the lack of experience and imagination of the ways in which various treatment courses may affect their life. In contrast, clinically-experienced physicians have the opportunity to develop vast experience with the kinds of social and moral concerns that may surface within their medical field of expertise and, thus, may offer some advice and direction on possible dilemmas.

Unfortunately, as it appears, many medical professionals assess their patient's capacity to usefully manage information found online as rather low; more importantly, they see such active and independent inquiry as impeding the efficiency of the clinical encounter (Kukla 2007, 33). Whether 'efficiency' is the relevant metric to be measured and evaluated in shared decision-making seems highly questionable. The proliferation of decision aids (Cochrane Collaboration 2009), thus, from a broader perspective may also be seen as emerging out of a need and desire for measurement (Epstein 1990) and the transfer of responsibility onto individuals (Rockhill 2005; Rose 2006). While these perspectives are significant, this research is focused on questions of design and extended cognitive processes in this regard.

Analysis of a decision aid

In analysing the design of the NHS decision aid on ‘Deciding whether to have prostate specific antigen (PSA) test’ (BMJ Group),⁸⁰ a number of issues and questions can be identified which are relevant for the practice-based research discussed in chapter four, five and six. To start with, the DA is a five page-long document, roughly structured into an introductory page, three pages of decision support regarding the PSA test, and one final page providing broader questions potentially useful to consider when deciding about the test as well as how to get access for further support.

On page one, the intended audience is stipulated as well as the suggestion that whether one should have a test or not depends on a variety of factors, such as “what you think you would do if you had a result that showed a raised chance of prostate cancer” (ibid.). As I will argue in my analysis here, unfortunately, it is precisely this aspect which seems unsupported in the decision aid. The structure of the actual decision support section (pages 2 to 4) is relatively simple. There are three columns: the first one lists the questions people may have, such as “What is the choice?” or “What is the effect on how long you live?”; the second one answers these questions from the perspective of having a PSA test; the third one does so from the perspective of not having the PSA test.

While the document appears well structured, the information provided seems a lot to understand and absorb. More technically, it seems very disturbing and confusing to me that the decision aid on page one clearly stipulates that PSA test cannot be used to diagnose prostate cancer with certainty, and then in the decision support section goes on to explain that “The PSA test can suggest you have cancer when you don’t (overdiagnose cancer).” Obviously, this is contradicting, wrong and potentially misleading. From looking at the information provided in columns two and three for a specific question, it stands out that oftentimes these texts are almost identical, with the only difference that in column three they are formulated in inverse terms. For example, with regard to the question “What is the effect on chances of being diagnosed earlier,” the answer in column two is “Prostate cancer is diagnosed 6-8 years **earlier** in men **who have** a PSA test,” whereas in column three the answer is “Prostate cancer is diagnosed 6-8 years **later** in men **who don’t have** a PSA test.”

The meaningfulness of these statements is partly conveyed in the following section concerning the question of “What is the effect of being

⁸⁰ See also Appendix 1.

diagnosed and treated early?” The response in column two and three is that it is unclear whether early diagnosis and treatment has an effect on length of life. Importantly, what is left out is supporting deliberation on the question stated as relevant on the first page, i.e. “What you think you would do if you had a result that showed a raised chance of prostate cancer.” For example, there is no mention of what it means to be diagnosed with prostate cancer and how that affects people in their lives, or, in Mol and Law’s (2004) terms, doing a disease rather than having a disease. With this I refer to supporting an imagination and sense of what it means to be diagnosed with prostate cancer and, thus, being able to deliberate the implications on a person’s own life. With reference to Kukla’s (2005) more nuanced notion of patient autonomy, I argue that the NHS PSA decision aid fails to achieve this as it does not include any support for preference construction or deliberation despite mentioning it.

On the final page of the DA, for example, they repeat that choosing the best option “means considering how consequences of each option will affect your life,” yet these consequences are hardly spelled out in the decision support section. For example, with reference to the previously mentioned aspect of being diagnosed 6-8 years earlier when having a PSA test, the DA does not mention the consequence that some people may suffer from iatrogenic effects, such as anxiety, following prostate cancer diagnosis, that is to say the lived experience of such a diagnosis.

In summary, it may be said that the NHS PSA decision aid is more about how a PSA test affects physical bodies and less about how the consequences of having or not having a test affect people’s lives. Following my argument about the importance of the preference construction process in this chapter, I will speculate in chapter four, five and six that there is substantial potential for redesigning this aspect of decision support. Furthermore, the DA is written in a rather dull style, not dissimilar to an insurance policy. Whether such a use of language is supporting people to meaningfully transfer the information provided to their personal and social life will be considered in relation to extended cognitive processes in chapter four.

Conclusion

In this chapter I have discussed the development of shared decision-making as a result of identifying small-area variations in medical practice that were not explainable by patient choice as well as forces contesting the paternalistic power of doctors over patients. Furthermore, I have analysed the design and evaluation of decision aids as one of the tools based on which shared decision

making is promoted. Based on the analysis, I have identified and discussed the ways in which the presentation of epidemiological risk information has implications for individual sense- and decision-making. As I argue, the question of how to design for uncertainty is an important and unresolved one. Specifically, I have highlighted that, while improving readability, the suggested and dominant style of using natural frequencies obscures deeper uncertainties inherent in prostate cancer screening decision-making. Within the context of decision-support for cancer screening, this raises the question of how they can be explored more productively together instead of mutually exclusive.

Furthermore, I have discussed and problematised patient autonomy with reference to preference-sensitive care situations, such as cancer screening, and suggested the need to reconsider the medical and social environment as relevant elements in the process of constructing preferences and values. Unfortunately, little research apart from controlled experiments exists concerning such preference-construction processes, in particular concerning the role experienced physicians may play in this. The following chapter will consolidate the critical review of the literature in chapters one to three and prospectively reflect on the issues identified from the perspective of and potential for redesigning decision support.

Chapter 4

Prospective reflection on issues in prostate cancer screening decision support

“There is nothing to fear (it may seem) but the probabilities themselves.”
(Hacking 1990, 5)

Introduction

This research project is interested in a politics and practice of interactivity in preventive healthcare decision-making by non-expert individuals. While shared decision-making (SDM) with its origin and motivation to identify, address and support preference-sensitive care situations sounded promising, as I have demonstrated in the previous chapters there are many issues and uncertainties involved in SDM for prostate cancer screening. Taking this as a starting point, this practice-based research is inspired by Mol and Law's (2004) suggestion to privilege action over knowledge. Following this, and more specifically, this research is interested in what such an approach would look like for SDM within the context of prostate cancer screening, and which theoretical and practical moves we would have to do in following such lines of inquiry.

My fundamental hypothesis is that shared medical decision-making within the context of prostate cancer screening could be productively supported by an exploratory model of interaction. I will argue that such a model of interaction is better suited to circumstances of sense- and decision-making under uncertainty, and, thus, to the aspirations of a patient-centred care seeking to deeply respect patients as unique living beings and care for them on their own terms (Epstein and Street 2011). The underlying cognitive processes for such an exploratory interaction model will be conceptualised as distributed and performed through the environment instead of being based on a information-processing model of mind and cognition. Beyond a critical analysis of the underlying interaction and cognition assumptions of current decision aids within SDM, the core contribution of this practice-based research is an examination of how we could start and go about designing for such exploratory interaction with extended cognition in mind within the specified context of prostate cancer screening.

The chapter is structured as follows: the first part will reflect on the issues in medical epistemology and decision-support as identified in the

previous chapters from the perspective of ontological politics. I will argue that the contemporary design approach to decision support is largely based on an information-retrieval model of interaction which tends to obscure the ontologically political issues underlying the decisional problem in preventive cancer screening. I will discuss the ways such a design approach is particularly problematic within the context of preference-sensitive care and the recently identified overdiagnosis and overtreatment of various cancers, ultimately failing to support the goals of SDM. In the second part of the chapter I will introduce the extended mind thesis and distributed cognition based on which this practice-based research will investigate the productive potential of an exploratory model of interaction to support shared decision-making.

Ontological politics in medical epistemology and decision-support

As I have discussed in the previous chapters, there are a number of issues contributing to the uncertainty involved in decision-making about prostate cancer screening. Among them is the epidemiological nature of data and risk factors, their applicability and usefulness for individual decision-making as well as the novelty of the decision situation and subsequent absence of established values and experience with it. Furthermore, the validity issues with EBM itself due to RCT's being performed with overly healthy participants not representative of the complete population, or the regional variances in care, are in and by themselves problematic within the context of SDM.

The epistemological assumptions underlying evidence-based medicine (chapter two) as well as the design of decision aids for supporting shared decision-making (chapter three) can be productively viewed through the lens of ontological politics. Such a perspective allows us to problematise and summarise key aspects which are more or less directly materialised in the design of contemporary decision aids and which this practice-based research aims to explore and start redesigning. Before venturing into reflecting on the identified issues from this perspective, I will start by introducing and discussing ontological politics as a term and its role for the methodological framing of this research.

Ontological politics

The contemporary use of the term 'ontology' is twofold. On the one hand, it is used in philosophical discourse; on the other hand, it is used in developing

and describing information structures in the area of software development. In philosophy, ontology refers to the “conditions of possibility we live with” and “what belongs to the real” (Mol 1999, 75) by which is meant an inquiry into the nature of knowledge of what exists and how these entities can be organised and categorised into hierarchies and subdivisions according to definitions of similarities and differences.

In software development and the computer and information sciences more broadly, the use and meaning of the term ‘ontology’ is as a descriptor of knowledge, or as Shirky puts it “an explicit specification of a conceptualization” (2005). In other words, it is used to conceptually describe sets of data, or data about data, also known as metadata. While analytical philosophy is stringent in its clarification of what has ontological quality, ontologies in the computer sciences are less rigorous in their epistemic endeavour and, as a consequence, have developed vastly, so much so in fact that they have, according to some, significantly increased the problem they were meant to solve (Smith et al. 2005).

The question of ontology in the medical field is particularly interesting. Regarding the philosophical use of the term, an example within the medical context would be that of a disease, such as cancer, its symptoms and pathologies. Diseases in medical science are classified in a nosology. As mentioned earlier, there are different nosologies, and diseases can be classified - among other aspects - according to their cause (etiology), the mechanisms by which they are caused (pathogenesis), or by symptoms a patient experiences.

However, issues with the meaning of the word ‘cancer’ and its continued validity and applicability have recently surfaced. Whereas according to popular belief cancer is typically closely linked to or expected to lead to certain death, recent research on cancer indicates that it “contains a whole range of diseases” (Esserman et al. 2013; Welch and Black 2010). Essentially, the problem is that screening detects many indolent lesions that may never require clinical attention and, thus, come to be known by the person who has them. Nevertheless, such lesions continue to be called cancer and the general sociocultural interpretation of such non- or not-yet malignant diseases has been strong for some time (Sontag 1977). On the basis of research on the dramatic levels of overdiagnosis in cancer (Welch and Black 2010), Esserman and Thompson (2010) go on to argue that a new definition of cancer is needed which makes a diagnostic, prognostic and explicit linguistic difference between malignant and indolent cancer.⁸¹

⁸¹ Their view has attracted its own critics: Lochlann argues that some of the arguments made by Essermann et al. are problematic due to their reliance on population data for retrospective

Mol (1999) refers to John Law inventing the term 'ontological politics' in his 1997 book *Aircraft Stories: Decentering the Object in Technoscience*.⁸² However, for our purposes regarding the medical context, Mol's own work is a more relevant source (1998, 144; 1999). Combining the terms 'ontology' and 'politics' has the intended effect of highlighting that the conditions by which we recognise and navigate reality are not pre-given and, thus, do not independently exist a priori of our interactions with such reality. Instead, as Mol (1999) stresses, reality and the conditions upon which it exists are configured within and through such practices. For her, "the term politics works to underline this active mode, this process of shaping, and the fact that its character is open and contested" (ibid., 75).

Mol, as well as others (Bowker and Star 1999), substantiates this argument well with numerous examples, such as the performance of multiple ontologies of *Atherosclerosis* (Mol 1998) or *Diabetes* (Mol 2008). In this work Mol cogently demonstrates the different ways in which medical professionals as well as patients do these illnesses in situ, and that such practices, in turn, have implications for the conditions of possibility. For example, with regard to diabetes Mol and Law (2004, 4) illustrate with their ethnographic work that doing diabetes or, as they suggest, "action is privileged over knowledge," may mean a variety of practices for different people. In the case of diabetes, a tightly regulated blood-sugar level scheme typically results in fewer complications with eyesight and arterial obstruction in later stages of life. However, and simultaneously, such a scheme increases the risk of hypoglycaemia, which, in turn, may cause brain damage (ibid.).

From an ontological political perspective, this implies that "medical interventions hardly ever bring pure improvement, plus a few unfortunate 'side-effects;' instead, they introduce a shifting set of tensions" (ibid., 58). Making sense of such a shifting set of tensions may well mean very different things in the lives of different individuals. As we shall see in the following section, such ontological politics have had implications (Wennberg 1982) and, as I will argue, continue to have implications which affect patients in less than desirable ways. Returning to the topics of evidence-based medicine and shared decision-making, I will now start summarising the analytical arguments in light of an ontological political perspective.

studies which is at risk of being unable for uncontrollable influences, hence of little comparative epidemiological value (Lochlan 2013, 157).

⁸² In following technological projects, Law (1997) demonstrates the complexities inherent and, more importantly, the ways in which these interact with each other. Thus, ontology is shaped by the way in which we interact with it in our daily practices, hence best not understood as an a priori and given truth.

Evidence-based medicine

Based on my analysis of medical epistemology in chapter two, various mechanisms in the production, distribution and application of medical knowledge can be viewed and prospectively summarised from an ontologically political perspective. A first fundamental change with ontologically political relevance occurred in the transition from bedside to hospital medicine in the western world, broadly speaking. As has been argued by Mol and Law (2004) and analysed earlier by Foucault (1973), while in bedside medicine doctors would make inferences from the sick person's experiential descriptions, in hospital medicine the sick person was muted and deemed incapable of knowing about his or her ailments.

This was facilitated by developments in the 19th century when doctors were trained in specialised and narrow fields of the body as well as equipped with a number of -scopes invented to further support the clinical gaze beyond the skin of the sick person. An implication of this is that some people today do not dare request further information from or even disagree with medical professionals for fear of being medically badly treated afterwards (Kukla 2007, 33). This is obviously rather problematic for the goals of SDM and in situations of preference-sensitive care, given that patient autonomy and, thus, a deep understanding of benefits and risks is at the heart of SDM.

A further methodological shift followed early in the 20th century with the introduction of epidemiological methods. As I have analysed in detail in chapter two, the principal aim is to identify and preventively manage potential pathology based on physiologically measurable indicators which are statistically linked to probable future pathology. Here we have a number of ontologically political moves:

Firstly, I argue that the assumption of the suitability of using postdictive⁸³ population data (Politi et al. 2007) as a reliable information source for individual decision-making in screening contexts is such an ontologically political move. This is because the act of framing epidemiological risk data and probabilities of a sample population as relevant for individual decision-making rather than public health management only is setting the reference to which an individual's risk understanding and decision-making is suggested to be taken. This is despite the inability to know any one individual's health outcome (Esposito

⁸³ I use postdictive here instead of predictive because, in effect, such data explains past patterns of incidence with regard to a reference population. Assuming a dynamic model of health and pathology (Canguilhem 1994; Hacking 2006), it is unclear whether and to what degree such data can predict outcomes in the future (Politi et al. 2007).

2007; Politi et al. 2007, 683). As Gigerenzer indicates (2002a, 4), the use of such quantitative and probabilistic information can lead to a whole host of possible misinterpretations. Recent research claims that the best way such quantitative risk information is presented without causing misunderstandings is as natural frequencies.⁸⁴ Nevertheless, the core problem concerning the representational character of such information for individual decision-making remains or, as I would argue, is even strengthened; this happens by generating the illusion of being able to control the future while being unable to prevent alternative courses of one's health and the subsequent need to account for one's previous action and decisions (Esposito 2007).

Secondly, and as a precursor to the use of epidemiological risk factors, as has recently been identified such samples taken for RCT's are not always very representative of general populations (Goldacre 2012, 110). The implication of this is that the epidemiological risk-factor information provides a somewhat false sense of security, climbing probability trees with preventive treatment while the natural course of pathology might remain unaffected (Greenhalgh 2014, 4).

Lastly, the use of such information is problematic because of the flawed ways in which EBM seems generated, including selective reporting and publication bias. These issues have rather direct implications for the conditions of possibility and shaping of realities. This is so because such evidence lays the foundations for what is considered valid knowledge and, thus, becomes available for the application and distribution of EBM in shared decision-making with decision aids and decision support, to which I turn now.

Shared decision-making and decision support

Following my analysis in chapter three, it could be argued that SDM emerged with an (implicit) motivation to counteract some of the ontologically political issues in medical epistemology and practice. An illustrative example of this is the problem of unwarranted small-area variation in healthcare delivery which cannot be satisfactorily explained by patient preferences or intrinsic pathology factors as documented by the Dartmouth Atlas (Wennberg 1982). This resulted in the differentiation between preference-sensitive care and effective care. To briefly recapitulate, the former are healthcare situations in which several treatment options are considered to have equivalent health outcomes when measured in terms of health economics indices, such as life-expectancy and quality of life months, with the treatments, however, affecting patients

⁸⁴ See discussion in previous chapter.

differently due to their side-effects. The latter are healthcare situations where a single best treatment has been evidenced (ibid.). SDM was initiated with regards to the need of involving patients in preference-sensitive care situations. The goal of such an involvement is to have patients fully informed about diagnosis and treatment options, including probable risks and benefits thereof.

However, despite the best of SDM's medical and moral intentions, I argue that there are a number of strong limitations in its contemporary design approach. Firstly, on a conceptual level, the design of the interaction with decision aids may be argued to rely on an information-retrieval model of interaction as analysed and discussed in chapter one. In short, the assumption is that a person has a precise and articulable information need, for example whether or not to take a PSA test for prostate cancer screening, which can then be directly and unambiguously responded to and upon which the person is satisfactorily equipped to take the desired decision.

Yet, as discussed above, the decision whether or not to pursue prostate cancer screening exhibits substantial uncertainties which go well beyond the question of whether or not to take a PSA test. To recapitulate, among the uncertainties are questions of representativeness in the sample population, whether or not pathologically causal forces remain stable, unknown lead time to pathology as well as the notoriously low discriminatory predictive value at the level of individuals. Furthermore, the implications of a decision to screen materialise over a long time-frame, in the case of prostate cancer typically 20-25 years, and concern personal, social as well as professional areas of a person's life. It is reasonable to expect that, as people engage in screening and live through the experiences of such things as having numerous false-positive PSA tests followed by repeated biopsies with negative results (see chapter seven), the chance that their values and perspective concerning the merits of screening might change. Thus, the current model of interaction may be too myopic in envisioning the full problem space individuals experience and, thus, unable to address the complex context and decisional problem at hand. For example, the NHS PSA test decision aid does not include aspects of the uncertainties involved in how to deal with the potential of a false-positive or false-negative PSA test result, apart from mentioning the possibility and probability of such an outcome. It does not offer decision support in how to grasp the full extent of its potential ramifications and how to tackle this potential or actual problem,⁸⁵ such as in the case of someone having already

⁸⁵ See also UK database for uncertainties of treatment effects, <http://www.library.nhs.uk/duets/SearchResults.aspx?catID=14666>; ac. 29.06.2014.

had a PSA test and negotiating how to handle the test result.

Furthermore, the information provided in the decision aid (DA) is based on the assumption that a user of a DA fits within the same epidemiological groups. It seems that the aim of articulating the relevance of such statistical information as “big groups of men” (BMJ Group) is to generate trust rather than supporting a critical examination concerning its representational character. Again, as I have discussed above, given the mechanisms by which such epidemiological risk information is generated, whether a specific person and user of a DA is statistically well represented is unknown. The DA does not provide any way to assess this statistical match and, thus, “requires a leap of faith” (Politi et al. 2007) from anyone using such information.

Secondly, the IR-model comparison seems also to *confirm* the apparent decision- and cognitive-process assumptions underlying the design approach of decision aids. The design of contemporary decision aids is usually focused on aggregating relevant knowledge and making it available in a structured, consolidated and supposedly easy to apprehend form and language. In terms of the decision process, the use-assumptions seem to be that: a) all the relevant bits of information can be aggregated and provided; b) this information will meet a (previously) set idea of what qualifies as a significant risk; c) the information processing is straightforward and a decision can be taken quickly, unproblematically and without future regret.

To recall, meta-analysis by the Cochrane Collaboration (2009) on the effectiveness of DA's indicates that there is a lack of satisfaction with the decision-making process. This may be because such a decision-support design approach, particularly within the context of treatment options that may have similar medical outcomes, ignores patient preferences and runs the risk of a silent misdiagnosis. Thus, as Mulley et al. have argued (2012), patient preferences matter precisely because, while the health outcome of a variety of available treatments is largely the same, the side-effects and uncertain outcomes of the various treatments have very different effects on a person's life. The silent misdiagnosis in this case is that of a wrong patient preference diagnosis, which, as they argue, typically goes unnoticed today (*ibid.*). As research has shown (*ibid.*), medical professionals tend to think of themselves as better at diagnosing patient preferences than they really are. In the case of benign prostate disease, research has shown that surgeons overestimate patients' preference to ameliorate urinary symptoms, where in actual fact patients prefer to avoid sexual dysfunction, which is a typical implication of surgical treatment (*ibid.*). Further such examples are plentiful (Moynihan et al. 2012). Preference diagnosis is defined as

a doctor's inference of what a patient would choose if he or she were fully informed. It is an inference because no patient – save perhaps the patient who is also a doctor and world-renowned specialist in the very disease with which he or she is afflicted – is fully informed. Preference diagnosis, like medical diagnosis, is often a best estimate based on imperfect information.

(Mulley et al. 2012)

While the above is a partly useful definition, I would complement it by explicitly recognising the various uncertainties regarding one's personal, social and professional life that typically co-exist when a person faces a novel medical situation. The difficulty of being well informed, being able to make sense of the information and deciding in accordance with one's personal preferences and values precisely stems from the novelty of the situation and the subsequent absence of a deliberated and established relevant set of values along which to orientate oneself (Kukla 2007).

By failing to articulate the underlying uncertainties discussed above more explicitly, the design of the DA seems to abstain from engaging the patient in the preference-construction process, a necessary element for coping with such uncertainties and best deliberating and developing values and preferences. In turn, such a design practice has immediate ontological relevance concerning patient autonomy, precisely because of the underlying uncertainties involved and the ways these can play out in the screening patients lives' as we shall see next.

Prostate cancer screening qualifies as a preference-sensitive care situation because, as it stands, it remains unclear whether early diagnosis and treatment "makes a difference to how long you are likely to live" (BMJ Group). This is because prostate cancer often grows very slowly and, in many instances, may not generate problems during a man's lifetime; in other words, more men die with prostate cancer than because of it (Ablin 2010). On the other side, the chances of prostate cancer treatment, such as removing the entire prostate gland (radical prostatectomy), resulting in side-effects are relatively real, with 29% for incontinence and 79% for impotence 5 years after treatment (Health Dialog, 16).

A further key argument why it is particularly relevant to do preference diagnosis in the context of prostate cancer screening and, thus, to analyse and consider questions of ontological politics is found in the recently identified levels of overdiagnosis and overtreatment of various cancer types. According to Welch and Black (2010, 605), cancer overdiagnosis can be defined as the diagnosis of a 'cancer' "that would otherwise not go on to cause symptoms or

death.” More specifically, overdiagnosis may also occur if a cancer never progresses (or may even regress) following detection, or when the cancer progresses at a rate slow enough for a patient to die of other causes before the detected cancer causes health problems. In the case of prostate cancer, the magnitude of PSA-detected prostate cancers is estimated at 60% (ibid.). The same phenomenon is present in a number of other cancer areas (see below). The level of overdiagnosis is manifest as the gap between rates of new diagnoses compared to morbidity.

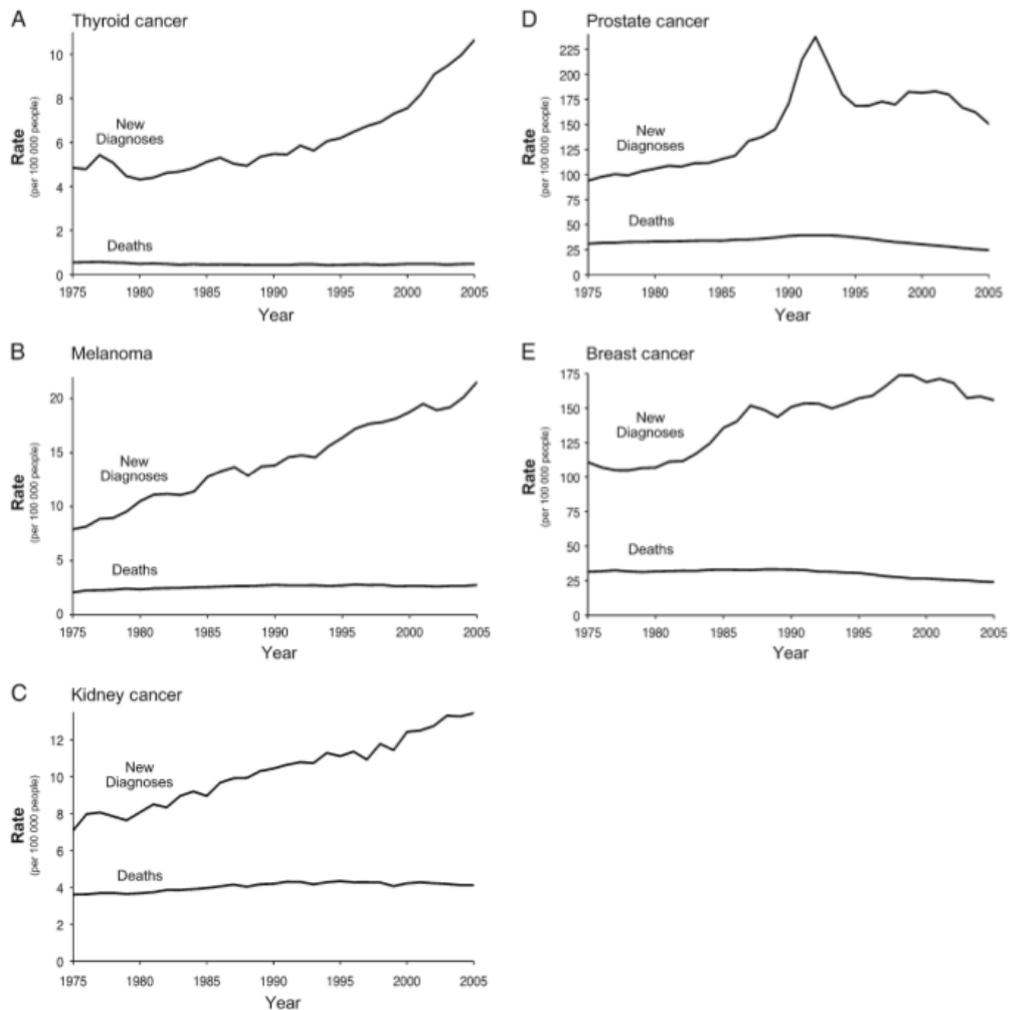


Figure 7. Rate of new diagnoses and death in five cancers in the Surveillance, Epidemiology, and End Results data from 1975 to 2005. **A)** Thyroid cancer. **B)** Melanoma. **C)** Kidney cancer. **D)** Prostate cancer. **E)** Breast cancer.

Figure 4.1: Overdiagnosis rates of cancers

Overdiagnosis and, consequently overtreatment can have substantial health implications for patients. As just discussed above, in the case of prostate cancer side-effects of treatment regularly include incontinence and impotence, not to speak of the psychological burden the fear of cancer puts on people and their families (Milton 1973). Early cancer detection, thus, may save

some lives while it disturbs and hurts many others.⁸⁶ As a result, taking an informed decision on early cancer detection requires an understanding of the benefits and harms involved from a patient's perspective. As Greenhalgh et al. (2014, 353) argue, "[W]e need to develop decision aids that support clinicians and patients to clarify the goals of care, raise and answer questions about the quality and completeness of evidence, and understand and contextualise estimates of benefit and harm."

Following my review of the assumptions embedded in the contemporary design of decision aids, generating an understanding of the benefits and harms involved in prostate cancer screening may be productively supported by explicitly including questions concerning the conditions of possibility as part of the design of the decision-making process. Specifically for this practice-based research, I will focus on the issues around using epidemiological risk information in decision aids as well as applying an exploratory interaction model to support preference construction. To this extent, I hypothesise that an exploratory model of interaction is better suited than an information-retrieval based model to the context and challenges of shared decision-making for prostate cancer screening. The former allows for a more open-ended design approach as well as hopefully supporting a more nuanced and complex understanding of patient autonomy and self-determination which seems adequate in the context of long-term and continuous screening decision situations (Kukla 2005). The key questions, thus, become: how to design for shared decision-making with such goals in mind? What can we learn about engaging with such ontologically political issues in healthcare information interaction and decision-making from a design process perspective?

Beyond a critical analysis of the underlying interaction and cognition assumptions of decision aids within SDM, the core contribution of this research is an examination of how we could go about designing for such exploratory interaction within the specific context of cancer screening sense- and decision-making. In exploring how to address the selected aspects of ontological politics in designing for SDM most productively, this practice-based research takes a different route by opening up the question of what can productively be considered as mind and, thus, involved in cognitive processes. I do so by assuming that, in some circumstances, mind and cognitive processes can be thought of as made up of things outside of the skin and skull, involving the social and artifactual environment. The approach I am following in this regard is that of extended mind and distributed cognition.

In part this move is inspired by Mol and Law's (2004) suggestion to

⁸⁶ See also chapter two.

privilege action over knowledge as discussed above with regards to the lived practice of doing diabetes. For them, the action/knowledge distinction is a way to “shift the grounds on which questions about the reality of bodies may be posed” (ibid., 45). They exemplify this by identifying the shifting set of tensions this may implicate for some, as discussed above with regard to diabetes and the benefits and risks of a tight regulation of blood sugar levels. Without negating knowledge, their perspective is more interested in the implications of medicine for the lived practice of being a patient and what we might learn from this. Such an approach, they hope, might help to integrate the role and perspective of patients back into medical practice, and allow them to account for the effects of medicine on lived experience rather than just how it affects their bodies (ibid., 58). Considering enacted bodies, as suggested by Mol and Law, is thought relevant for framing the extended mind cognitive approach and the design approach taken in this research. Moving beyond Mol and Law’s suggested distinction between action and knowledge, the approach I am following is interested in the epistemic insights following from and through embodied action over time.

Extended Mind Thesis

In contrast to the conceptualisation of human cognition as developed in the preceding chapter, this section will introduce the Extended Mind Thesis [EMT] as an alternative perspective on human cognitive processes. In light of the criticisms of classical ideas of human cognition, I will argue there are convincing arguments that EMT allows for a productive engagement with the issues identified in exploratory health information interaction discussed above. In particular, I will argue that the approach of EMT is advantageous for rethinking and redesigning for interactive (cognitive) processes because it recognises wider kinds of artifactual and social ecologies that may productively support such processes.

This section will develop as follows: firstly, I will introduce and explicate EMT. This will serve a dual purpose. Firstly, a detailed consideration will explain the different theoretical and empirical insights that combine into EMT, rendering, on the other hand, its potential relevance more compelling with regards to some of the issues identified in the three preceding chapters. Secondly, by way of selected examples, I discuss and illustrate some of the more specific arguments of EMT. Thereby I identify the crucial elements in which it differs from the classic information-processing model of mind and, more importantly, demonstrate its particular relevance and potential for

exploratory health information interaction. Significantly, this section as well as this practice-based research more broadly is less concerned with the intricacies of philosophical and cognitive science arguments concerning ideas and questions of computationalism and representation in competing theories of human cognition and how EMT accords to them. Instead, I am interested in researching the ways in which taking an EMT perspective would affect the redesign of exploratory health information interaction.

Crucially, and in contrast to classic cognitive theories, in EMT cognition is not delimited by processes that occur within one's brain, but extends into and operates through the social and artifactual environment. This is because, as evidence suggests, such artifactual and social ecologies play a key role in cognitive processes as well as a determinant role in human development and evolution more fundamentally. The thesis of extended mind for Clark and Chalmers (1998), who first formulated EMT, is that "when parts of the environment are coupled to the brain in the right way, they become parts of the mind" (Clark 2008, x). The role of the artifactual environment in cognition for EMT in its simplest form may be usefully illustrated by considering a number of examples.

Firstly, the specific case of patients suffering from Alzheimer's disease: a typical symptom of the disease is a progressive mental deterioration leading to dementia and memory loss, at first concerning short-term events only and over time also long-term ones. A case in point was a group of Alzheimer's disease patients living in inner city St. Louis, USA (Clark 2003, 139). What was surprising is that these patients were able to successfully live independently despite the fact that they performed rather poorly on standard psychological tests evaluating their cognitive capabilities as well as compared to other patients at such a stage of disease progression. What seemed to have made this independent living possible was the fact that their homes were equipped with numerous (cognitive) scaffolds that supported memory for various, everyday life practical as well as social situations. Among the scaffolds were

message centers where they stored notes about what to do and when; photos of family and friends complete with indications of names and relationships; labels and pictures on doors; 'memory books' to record new events, meetings, and plans; and 'open-storage' strategies in which crucial items, such as pots, pans, and checkbooks are always kept in plain view, not locked away in drawers.

(ibid., 140)

While one could be quick to see such external scaffoldings as mere

external memory, Clark and Chalmers provide a more subtle argument. For them, such arranged mental notes regarding, for example, family relationships have been personally generated and, thus, endorsed. In this way, in the case of Alzheimer's patients, they seem to not only work as external memory, which they do, but also to generate the necessary condition for their acceptability. From personal experience the argument holds equally true in the case of making innumerable reading notes during extended periods of PhD research. Rereading such notes after a significant pause sometimes leaves one in wonder that one has thought a specific thought. In some cases, at least, if it were not for the recognition of one's personal hand-writing, one's doubt of being the author of it would seem almost insurmountable.

Secondly, with regards to the role of language, researchers conducted a study involving prelinguistic infants who were shown the position of food or a toy in a room. They were, then, disoriented and required to try and find the previously shown food item or toy. Importantly, the room was constructed in such a way that the "location [of the item] was uniquely determinable only by remembering conjoined cues concerning the colour of the wall and its geometry (e.g. the toy might be hidden in the corner between the long wall and the short blue wall)" (Clark 2008, 48) Thus, either geometric or colour cues independently would not suffice to locate the items. What the researchers found was that prelinguistic infants would only make use of the geometric information, "searching randomly in each of the two geometrically indistinguishable sites" (ibid.), yet when conducting the experiment with adults and linguistically capable children, participants combined geometric and colour cues to successfully solve the problem. Furthermore, successful combination of the cues was independent of aspects such as intelligence or developmental stage. Instead, it was only children able to combine spatial and colour descriptions via use of language who could successfully solve the problem (ibid., 49).

Lastly, with regards to the role of language, a study was conducted with Russian-English bilinguals who were trained in sums of two-digit numbers presented as words in one of the languages. The test then compared participants' ability to approximately and exactly select their answers from a given choice. What the researchers found was that test candidates were successful in selecting the correct approximate answers independently of the language they were trained in. However, the ability to correctly select the exact answers depended on the language they were trained in and was much slower in the untrained language (ibid., 51). Interestingly, the way the researchers make sense of these findings distinguishes three distinct cognitive abilities: firstly, a basic biological ability to distinguish small quantities, such as "1-ness,

2-ness, 3-ness and more-than-that-ness” (ibid.); secondly, a basic biological ability to distinguish magnitudes, such as an “array of 8 dots from arrays of 16 dots but not from more closely matched arrays” (ibid.); thirdly, the learned use (and thus not biological capacity) of number words of a language and, more importantly, that these number words refer to specific and distinct quantities (ibid.). What these three cognitive abilities add up to is the capacity to acknowledge that “the number word 98 names a unique quantity between 97 and 99” (ibid.).

This last case is particularly interesting and will be relevant in later sections when discussing the use of precise quantitative probabilities in relation to medical decision aids. This seems particularly so because it is difficult to imagine that anyone can have a clear image of, for example, a 98-ness as compared to a 2-ness; yet, as a consequence of our learned capacity to individuate distinct number words, we see it as a precise number in a linear and robust sequence of quantities. This is especially problematic in light of the various contingent factors constituent in generating probabilistic medical evidence, as discussed in chapter two.

The exactitude of such a probabilistic figure holds with respect to the methodological assumptions, such as an infinite number of experiments, a random distribution of error and with regards to a sample of a reference population but not necessarily for the decision-making individual patient. This is because, by definition, from an individual, biological body perspective, there is no such thing as a probability to get cancer: one either develops it or not. Probabilities work the other way around: they are a systematic collection of individual, actual biological realities expressed in a probabilistic manner as based on the presumption that there will be no underlying biological change in the constitution of individuals and the population. As Canguilhem has cogently argued (1994), this may be a very myopic perspective.

The examples aim to illustrate the use of external scaffoldings as well as the role of language as a tool for solving various cognitive problems in situ. Nevertheless, Clark (1998, 179) rightly raises the question of whether these indicative studies on the role of the artifactual and sociocultural (language) ecologies are also able to extend the role of EMT with regards to making sense of more complex social and cultural interactions, such as “[political] voting, consumer choice, planning a two-week vacation, running a country, and so on.” To this list I would personally add making sense of complex medical information from various sources, or taking medical decisions regarding treatment or therapy. Clark posits that EMT is equally applicable and, in such circumstances, the external structures are more of a symbolic and social-institutional character (ibid.). In these more advanced or complex cases of

human cognition, he suggests that the role of public language is particularly relevant for

social coordination and individual thought. The idea, in short, is that advanced cognition depends crucially on our abilities to dissipate reasoning: to diffuse achieved knowledge and practical wisdom through complex social structures, and to reduce the loads on individual brains by locating those brains in complex webs of linguistic, social, political, and institutional constraints.

(ibid.)

To that end, and following EMT, cognition may in some cases be better described as (more or less significantly) influenced and determined by embodied and socially situated circumstances.⁸⁷ The following section will discuss various sources and insights from related fields from which EMT emerged and the ways in which it differentiates itself from these other cognitive science theories.

Venture points for Extended Mind Thesis and related fields

EMT builds on and has evolved from insights from a number of different fields. The following discussion aims to identify and explicate its most relevant sources of inspiration. The detailed consideration will explain the different theoretical and empirical insights that combine into EMT. This will render its relevance more compelling for beginning to address the issues with health information interaction I have identified in the preceding chapters.

EMT may be said to have a short history and a long becoming. As Clark notes in the chapter titled 'Groundings' in his first book on EMT (1998), the idea of mind as deeply linked through action of bodies in the world is at the heart of Heidegger's *Being and Time* (1927). One important aspect of Heidegger's phenomenology was the notion of 'Dasein,' which he has introduced to overcome Husserl's earlier mentalistic concept of intentionality. For him 'Dasein,' or 'being-in-the-world,' was the essence of how we encounter the world and make sense of it, that is intentionality being inseparable from being and doing. He illustrates his concept with the example of the hammer. We sometimes act through technology (ready-to-hand) and sometimes are

⁸⁷ Whether or not one subscribes to the cognitive load reduction part of the argument is open to discussion. While such discussions are not at the core of this thesis, I do not see the need for such a dependency to argue for the plausibility of extended and distributed cognition.

conscious of the tool (present-to-hand) when we need to readjust a tool. For him the tool comes into being through the transition of ready-to-hand to present-to-hand, thus a process of meaning making as practical and purposeful. For Paul Dourish (2004, 138), an interaction design researcher, “intentionality sets up a relationship between embodied action and meaning.”

Such considerations are further illustrated by the more recent work of Alva Noë (2009), a philosopher and neuroscientist, who has reinforced the argument for an understanding of cognition as embodied and enacted. Building on the work of Maturana and Varela (1987) and others (Dreyfus 1979), he cogently demonstrates how the sensorimotor human apparatus is intricately involved in the process of seeing. To do this, he refers, for example, to the work of Bach-y-Rita, a neuroscientist working on neuroplasticity, who has built an apparatus for sensory substitution that should enable blind people to ‘see’ again (Noë 2009).

Essentially, Bach-y-Rita believed that the eyes were a channel for getting visual information to the nervous system and it should be possible to achieve this using a different channel. Therefore, he built an apparatus with a camera that was linked to an array of vibrators that were placed on the thigh or abdomen of subjects. The camera was mounted on a blind person’s head or shoulder and, upon use, generated tactile sensations on the skin of the subject. The subject could then make judgments about the size and shape of rooms, objects within it or even reach out and pick up objects. In fact, subjects only needed a few hours to get used to the tactile sensations. Furthermore, what is interesting was that somatosensory touch areas of the brain were then giving rise to a visual experience rather than one of touch. Thus, use and action context generated perceptual plasticity.

The insights for embodied cognition can further be supported by considering the perspective of developmental biology of cognition (Griffith and Stotz 2000, 37). Recent evolutionary epistemology rejects a nativist view on the evolution of cognition, which is predicated on the idea that mental capacities are innate. Such nativist perspectives hold that “[M]inds are constrained to develop in a particular way, and, once developed, they are constrained to reason in a particular way” (ibid., 30). In contrast, one strand of evolutionary epistemologists view constraints as enabling, instead of a mere reduction of possibilities (ibid., 31).

According to such a perspective on cognitive development, outcomes are not pre-programmed but derive from a cascade of interactions between organism and the environment. Among the empirical evidence advanced to support such arguments are experiments with infants in which certain parameters of a system are selectively changed. For example, coordinated

stepping, while present from early infancy, is usually suppressed due to the relation of the weight of limbs to muscular force. When such a physical constraint is removed, the walking behaviour can be activated much earlier (Thelen and Ulrich, 1991, 38 cited in Griffiths and Stotz 2000). As a consequence of such theoretical arguments and empirical evidence of a developmentalist approach, Griffiths and Stotz (2000) argue that constraints in development adopt a 'soft' role. 'Soft' in the sense of their nature as probabilistic, instead of deterministic, because contingent on a wide range of developmental resources. At the heart of such considerations, and in relation to EMT, is the insight that meaningful interaction with the world seems to profoundly rely on intentional interactivity facilitated by various means and channels of perception in action.

Further areas relevant to consider are 'epistemic action' from within the field of cognitive science and 'distributed cognition' as developed by Edwin Hutchins (1995a), a cognitive anthropologist. Epistemic actions, as Kirsh and Maglio (1994, 513) claim, are "actions performed to uncover information that is hidden or hard to compute mentally" as differentiated from pragmatic actions "performed to bring one physically closer to a goal." Kirsh and Maglio observed players of Tetris, an interactive video-game in which the player needs to arrange objects of various shapes in such ways as to fill in rows at the bottom of the screen. Whenever a row is fully filled, it disappears and makes space available. When rows cannot be fully filled, they will build up and, thus, allow less space to manoeuvre the falling objects. While the objects fall from the top of the screen, the player can either rotate or move them from left to right.

What Kirsh and Maglio observed was that "certain cognitive and perceptual problems are more quickly, easily, and reliably solved by performing actions in the world than by performing computational actions in the head alone" (ibid., 513). The authors' interpretation of such actions in the world is that they improve cognition. An exemplary epistemic action is the turning of objects to more easily identify their shape or moving to the far right to determine exact position for high-drops. From their study, the authors concluded that standard information-processing models of Tetris cognition are unable to explain many of the actions performed by the players and make them seem unmotivated and superfluous. Furthermore, they found that such "traditional accounts are limited because they regard action as having a single function: to change the world. By recognizing a second function of action - an epistemic function - we can explain many of the actions that a traditional model cannot" (ibid.).

Lastly, there exists the interesting research on external scaffoldings and

distributed cognition by anthropologist Hutchins. Hutchins is interested in the cognitive processes that occur in navigation, particularly large military vessels (1995a) and airplane cockpits (1995b). What emerges from research on such processes of cognition according to Hutchins are three kinds of distributions. Firstly, the observation that cognitive processes may be distributed across members of a social group. Secondly, cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structures. Thirdly, cognitive processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events (ibid.). The argument for causal coupling goes against much of mainstream grain of brains as representational machines endlessly busy creating internal models of the external world.

As this and other diverse research has come to suggest and support (Bach-y-Rita et al. 1969; Brooks 1991; Clark 1998, Maturana and Varela 1987; Noë 2009), cognition emerges out of a much more complex entanglement of internal and external processes involving perception, attention, memory as well as the material and cultural environment (Hutchins 2000, 177). Hayles (2012, 93) articulates the difference between embedded and extended cognition well: “Whereas the embedded approach emphasizes human cognition at the center of self-organizing systems that support it, the extended model tends to place the emphasis on the cognitive system as a whole and its enrolment of human cognition as part of it.” In embedded cognition, people use various objects as scaffoldings to support or extend memory and, thus, make possible more sophisticated thinking. Hutchins, an anthropologist, demonstrated such use of objects in relation to the complex tasks of navigation (Hutchins 1995b). Extended cognition, in contrast, places much less emphasis on the human mind and, thus, agency as the central controller in cognitive processes (Hayles 2012, 94).

As such, EMT as an approach also stands in sharp contrast to multiple other ideas and theories of cognition, such as, for example, exclusively genetically determined ideas of human development (Pinker 2003) and, as discussed earlier, brain-bound information-processing models of mind and rational choice theory. While rational choice theory itself is not a theory of cognition but more at home in the field of economics, it is based on cognition as brain-bound information processing and relevant to consider in relation to decision-making theories. Genetically deterministic ideas of human development will not be discussed further, but ideas of rational choice and the information-processing model of mind are more close to the topic of this research and worth illuminating.

Of particular interest at this stage is considering the cases and circumstances in which rational choice theory seems to work as well as, more importantly, where it fails to predict outcomes. Ironically, as several authors have noted (Satz and Ferejohn 1994), it appears that “neoclassical economic theory works best in situations in which individual rational choice has been severely limited by the quasi-evolutionary selection of constraining policies and institutional practices” (Clark 1998, 182). More specifically, what seems to happen is that such institutional structures have evolved in environments of strong competitive forces, for example, capital markets, which in turn result in firm-level strategies and policies with little leeway for individual beliefs and values (ibid.). Furthermore, in other environments, strong communities of practice have evolved that exhibit relatively set values and practices of how work is to be done (Wenger 1999).

Consequently, this also means that in light of this theory, cognitive processes are constituted by interactions with a complex ecology of human and artifactual elements. Maybe more importantly, such an “expanded notion of cognition should make us cautious about the kind of things we design and build. Institutions, training, forms of media, linguistic systems, and a multitude of other factors will all find new meaning in the light of this theory” (Fuller and Matos 2011, 10). This is precisely the issue Clark raises when he suggests that “[O]ur biological brains, in concert with these new technologies [search engines, digital networks] can thus grow into hybrid minds better able to understand the kinds of systems in which they themselves participate” (Clark 2003, 159).

The kinds of ways in which human cognition seems to operate through and participate in artifactual ecologies have also been recognised and come to be exploited more opportunistically in the field of behavioural economics. The field has been popularised by Thaler, an economist and behavioural scientist, and Sunstein, a legal scholar and behavioural economist. Interventions following this approach are based on the idea that the artifactual environment can be designed in such ways that it ‘nudges’ people to behave more likely in ways thought to be more beneficial for them than others. For example, healthy foods would be placed at the beginning of a long array of food-displays in a school canteen rather than at the end. This, it is believed by proponents of behavioural economics, will make it more likely for students to choose healthy foods than otherwise. Such an approach goes by the supposedly self-explanatory term ‘choice-architecture’ (Thaler and Sunstein 2008). Typically, evidence for the performance of such an approach is experimental. Indeed, as Gigerenzer, a psychologist, and Berg, an economist, argue, the evidence base is rather thin because, rather than researching how

people actually make decisions, it is only looking at what decisions they make and then generalise from such experimental evidence (Harford 2011).

From a more political perspective, a critique voiced is that 'nudge' interventions are seen as 'liberal paternalism' because they are designed and imposed top-down and, due to their nebulous presence, do not invite participation, reflection and, thus, learning and long-term behaviour change. However, it is precisely these aspects that need to be considered in order to tackle some of the more vexing problems, such as changing unhealthy diets and helping people stop smoking (Rowan 2011). As Esposito (2007) has noted, planning (based on probabilities) does not allow the refusal of learning from actual experience, which is particularly true when confronted with negative outcomes following a decision no matter how likely it was supposed to turn out another way. In turn, the preventive action of reducing one's statistical risk neither pre-empts a less desirable course of action nor prepares the person to confront and handle such a situation. At the core then, engaging with the uncertain character of some medical interventions, such as cancer screening, sooner or later requires to recognise and confront the shifting set of tensions introduced by such interventions.

Specifically problematic for this research is the anticipatory practice of surveillance-based medicine (Armstrong 1995; Topol 2012). Such active screening of the population is aimed at early identification and prevention of specific conditions, such as diabetes, heart diseases as well as various cancers, including prostate and breast cancer. Screening procedures particularly operate on the epidemiologically determined probable link established via a linear gradient smoothly connecting normality (being clinically healthy) with risk factors (with or without symptoms) and pathology as defined by measureable quantitative changes and corresponding probabilities.

From an individual and social perspective, risk factors being causally linked to probable later pathologies can easily be imagined to morally implicate necessary preventive action. Furthermore, as has been highlighted, the application of such epidemiological methods has also been extended with regards to the ability to affect and control the conditions of new life, as is the case with screening of pre-born children (Rose 2006). While being probabilistically at risk or actually ill, as defined by the averages of derived epidemiological data and linked by the measurement of physiological factors (such as blood glucose level or blood pressure), an individual may still be free from experiencing actual symptoms. Such situations with regards to diabetes have initially been labelled protodiabetic, chemical diabetic, latent diabetic, stress diabetic and pre-diabetic, and later been reclassified for clinical

diagnosis as chemical, mild or early diabetic (Greene 2007, 100, 106). Clearly, the labelling chosen suggests a causally determined pathological progression, with little concern for its probabilistic character. Thus, it seems easy to imagine an individual's perceived risk of potentially becoming ill in light of the weight of medical nomenclature and the causally argued probabilistic evidence.

From an extended-mind thesis perspective, the way in which medical evidence is defined, developed and distributed seems particularly problematic for the following reasons: firstly, because of the ways in which the evidence is presented as factual with natural frequencies⁸⁸ rather than probabilistically and, thus, contingent on the various assumptions and factors based on which it was derived. The display of a single quantitative (probability) figure as representative and useful for individual decision-making about a clinical test, such as that found in decision aids concerning taking a PSA test (BMJ Group),⁸⁹ seems highly problematic. This is precisely due to the potentially significant implications that may follow from such a simple blood test that may follow a false-positive test result. These implications include mental-stress with test result, errors in test result (false positive and false negative), indicated need for a biopsy to confirm PSA test results and, thus, the potentially serious physical and social side-effects following preventive or early treatment, such as incontinency and erectile dysfunction (ibid.). From reviewing decision aids, such quantitative information is increasingly given in the form of natural frequencies (Gigerenzer and Hoffrage 1995). For example, the BMJ Group decision aid for taking a PSA test states the following:

More men who have a PSA test find out they have prostate cancer than those who don't have a test. Between 6 and 7 in 100 men who have a PSA test are diagnosed with prostate cancer. Between 4 and 5 in 100 men who don't have a PSA test are diagnosed with prostate cancer.

(BMJ Group, 3)

While these relative figures may hold true statistically speaking across a certain sample population, the distinctive figure is not suitable to predict one's own outcome, as in fact the actual outcome for oneself at the end of the diagnostic process is a binary between having or not having prostate cancer. More likely it is serving as an approximation to assess the cost-benefit trade-off between, on the one hand, taking the test and risking potential side-effects,

⁸⁸ Such as "Ten out of every 1,000 women have breast cancer." See also chapter three.

⁸⁹ See also Appendix 1. This is a decision aid distributed by the NHS which is supposed to support individuals in deciding whether they should conduct a PSA test or not.

and, on the other hand, not taking the test and risking undetected prostate cancer and its potential implications.

From an EMT perspective, it seems rather questionable whether and how a non-statistical expert is able to make sense of the range of contingencies involved in such a simple quantitative probabilistic figure. One may wonder whether the presentation of a PSA test statistic as natural frequencies figure, instead of a more contingent framing of it, favours its acceptance and somehow suppresses a more careful and deliberative reflection on its meaningfulness and situated applicability. In some way, it is as if quantitative probabilistic evidence can linguistically determine someone to be poised to be at risk or even ill. As Politi et al. point out (2007, 692), it is hitherto unclear “how patients respond to personalized risk estimated;” Rockhill (2004, 127) adds to this that “we are participating in open and unsystematic experimentation in these arenas.”

Secondly, probabilistically-derived risk-factor evidence is presented and framed as already causally linked to potential, and in some cases likely,⁹⁰ pathology at a later stage unless preventive measures are taken. As discussed above, in some cases, such as with ‘mild diabetes,’ even the nomenclature used incorporates such a deterministic logic. The question from a cognitive perspective then is how, despite such naming, people make sense of the probabilistic nature of such evidence in light of its linear, causal and pathological framing.

An interesting parallel might be drawn at this point: Merlin Donald, a psychologist and cognitive neuroscientist, in his exploratory investigations on the evolution of cognition and culture suggests that much of pre-Greek use of language was ‘mythic,’ i.e. exclusively used to record myths and complete theories. In contrast, the Greeks began using language to record and circulate “processes of thought and argument” (Clark 1998, 206). The notation of partial knowledge invited others to participate in developing the ideas further, criticising and complementing them. For Donald, the result of this was “much more than a symbolic invention, like the alphabet, or a specific external memory medium, such as improved paper or printing,” but rather “the process of externally encoded cognitive change and discovery,” as such language moves beyond being a medium for information communication and assumes a tool-like character (Donald 1991, 339).

From an EMT perspective, this is because such language opens up text as an object to reflect upon and, as Clark describes with regards to his own writing practice, that it is “these resources [that] enable us to pursue manipulations and juxtapositions of ideas and data that would quickly baffle

⁹⁰ Such as when the percentage probability of an event is larger than 50%.

the un-augmented brain“ (Clark 1998, 206). To be sure, there is nothing deterministic in either the pre-Greek or the Greek use of language itself; more importantly, it seems the culture developed by the Greeks around the use of textual language may be said to have facilitated, or as Donald puts it ‘invited,’ a more explicit cognitive reception and response to it.

Comparing such processes of language use as tool-use to the presentation of medical probabilistic facts begins to illustrate that, for the non-expert user, medical language may be said to display (yet) little possibility for explorative interaction and sense-making (literally) with it as a scaffold. This, I speculate, may be largely due to the difficulty in understanding this very specific and technical language, including the very assumptions and methods it is built upon. Such understanding is further undermined by the ways in which medical information is presented and is, thus, less amenable to a critical examination of its applicability. Comparing this to the previously discussed perspective of patient-driven research in the participatory medicine model, Scherzer (2012) maintains that “an important component of the patient contribution is to balance what evidence-based medicine or conventional medical wisdom recommends with what is possible, desirable and most acceptable for the individual patient.”⁹¹

Consequently, the current approach to designing for shared decision-making can, in the words of Merlin Donald, be said to be oriented along mythic trajectories rather than theoretical ones. This is because it conceals the panoply of representative assumptions it is built upon, which may simply not hold in an individual case. In turn, this concealment also occasions the potential for a mismatch between a hoped-for probable future and the actual, and unprepared for, pathological development. Furthermore, overdiagnosis and overtreatment may in this light also be seen as a consequence of, among other things, defensive medicine (Moynihan et al. 2012). Preference construction and diagnosis is a timely and necessary task, balancing and complementing a decision on probable futures with likely implications and their meaningfulness for individuals. Extended cognition, in this regard, is hypothesised as a productive approach in supporting such processes of preference construction. This is because it employs language as a means by which the exploration of values and preferences can be supported by way of recognising, questioning, juxtaposing and deriving them in spoken and written language. This, in turn, helps identifying shortcomings and novel insights into our own thinking about values and preferences.

⁹¹ Personal communication with Norman Scherzer, Executive Director Life Raft Group, 19. 04. 2012. The Life Raft Group is a foundation supporting research, patient support and advocacy for patients with gastrointestinal tumors.

With regard to the focus of the practice-based aspect of the present research, it is specifically relevant to consider the case of preventive screening for prostate cancer and to reflect on the effects that decision aids seem to have on the interpretation of probabilistic medical evidence, such as the diagnostic effectiveness of a PSA test. Prostate cancer has a relatively low mortality rate of 3% due to the fact that most of it is non-aggressive in nature. The consequence of that is that most men of those who have it die with prostate cancer rather than because of it (Ablyn 2010).

However, there are serious potential iatrogenic (typically negative consequences following medical intervention) implications from having either a biopsy [such as rising figures of bacterial infection which may lead to hospitalisation and in 9 out of 10,000 cases even to death (Gale 2011)], which is the only method based on which an actual diagnosis of prostate cancer can be made, or actual (preventive, early or overdiagnosed) prostate surgery or radiation treatment (potentially leading to incontinence and/or erectile dysfunction). Lastly, the reduced risk of dying from prostate cancer when taking a PSA test is 0.1%, or 1,000 men need to take the test to prevent 1 death (BMJ Group). Currently, 30 million (Ablyn 2010) men take a PSA test a year in the USA, which is approximately 54% of the male population older than 40 years of age.

From the evidence available it seems that the minimum we can say is that men in this case seem rather risk-averse concerning the long-term risk of developing cancer but willing to take various short-term risks of serious side-effects from cancer screening tests and prevention (over-)treatment. To some degree, I speculate, such a pattern might be said to be only a little surprising, as the factual evidence in the decision aid seems partially unbalanced. According to Lumen et al. (2012), diagnosis of prostate cancer happens six to eight years earlier in men who take a PSA test. Currently, it is unknown whether early diagnosis and treatment has an effect on expected lifetime (BMJ Group).

The above obviously leaves open the question of whether concern about prostate cancer is less about dying earlier, and more about the (potentially feared) implications for quality of life. The BMJ Group decision aid further points out that between 65 and 76 men in 100 (Heijnsdijk et al. 2012; Lane et al. 2010) with raised PSA-levels turn out not to have prostate cancer upon biopsy. Unfortunately, there is a lack of knowledge on the percentage of men who will suffer from the various potential iatrogenic effects of biopsies and/or cancer treatment (BMJ Group; Mulley et al. 2012). As research suggests, typically risk information is sparsely shared and discussed with patients (Politi et al. 2007, 690). When considering the evaluation of the performance of decision aids, fewer men seem to be willing to test for PSA

(O'Connor et al. 2009, 41). These effects, it may be speculated, trigger decision avoidance due to becoming aware of the uncertainties involved (Politi et al. 2007, 688).

The need for a more socially embedded understanding of medical language and knowledge has recently been recognised, as, for example, in suggestions made for a 'Health Knowledge Commons' (Lodor et al. 2013). The proposal for such a commons includes a platform for shared sense-making of health data and what such data actually means in people's everyday and social lives. This idea also leans on existing experience from patient-driven research platforms, such as 'Patients Like Me,' where people can record and exchange disease-specific information on their condition, experiences with treatments and therapy as well as resulting progress.⁹² In light of the recent evidence on levels of overdiagnosis and overtreatment in prostate cancer (Welch and Black 2010), such a move seems timely.

Summary of prospective reflection

In conclusion, this chapter has sought to introduce extended cognition as an alternative theory of human cognitive processes with particular relevance for shared medical decision-making processes. Using illustrative examples, I have tried to argue for its potentially productive role in informing research and design of complex health information interaction processes. As I have discussed, its attractiveness stems from a view on cognitive processes as enacted through the material and sociocultural environment, in which language assumes a particularly significant and productive role.

Furthermore, I have discussed the productive potential of EMT specifically with regards to addressing problematic issues in the design of decision aids. These issues include the use of quantitative probabilities; the IR model-based design approach which obscures more fundamental uncertainties underlying such risk factor information; the time-frame considered which ignores longer-term ramifications following false-positive test results or the realisation of an overdiagnosis; and the absence of supporting the construction of preferences and values as a source of decision guidance. Redesigning decision-making processes with the premises of EMT in mind is hypothesised to productively support making sense of probabilistic risk factor information as well as the preferences-construction process and, thus, shared

⁹² See also <http://www.patientslikeme.com/>. ac. 21.08.2014.

decision-making. In the following chapter, I discuss the selection of issues from the critical analysis which will be addressed and operationalized in the practice-based research. In chapter 6, I discuss the methodological approach and the design process.

Chapter 5

Focus of the practiced-based research and its operationalization

Introduction

The aim of this chapter is to establish a productive bridge and transfer between the critical analysis in chapters 1 to 4 and the practice-based design research in chapters 6 to 8. This is necessary and relevant for the following reasons. The critical analysis has looked at a variety of relevant topics as they relate to the interest of this research, that is the problem of how to start redesigning cancer screening decision-support as regards notions of risk and uncertainty. In order to provide a productive ground to inform and inspire the practice-based design, the critical analysis has required comprehensiveness and depth. Numerous facets of the problem of decision-support in cancer screening have been identified through the critical analysis. Given the early stage of this research within the larger problem context, the practice-based research will focus on a particular selection of these issues.

This chapter thus aims firstly to link the problems identified in the critical analysis as they relate to decision-support in cancer screening by discussing the informational health practice that follows from the critical analysis.

Specifically, my critical analysis finds that the current design of online search engine interactivity is strongly oriented towards an information-retrieval (IR) use scenario. The IR orientation has been further elaborated by way of algorithmic filtering of search results (Feuz et al. 2011). Such an approach however, further obscures the health information seeking process which leads me to argue that the exploratory nature of health information seeking interactions is badly supported by universal search engines today. Alternatively, an exploratory model of interaction is proposed as productive for guiding the design of health information interactivity, both online and offline.

Secondly, given the critique of probability theory that was established in the critical analysis, my position on how it relates to the practice-based design phase will be discussed. The proposed relevancy of an exploratory health information interaction approach becomes more apparent when considered within the context of making cancer screening decisions based on probabilistically derived risk information. Following my critical analysis of EBM, I argue that such risk information undermines the fundamental inherent uncertainties at the individual decision-making level. Such uncertainties are further obscured by risk representations based on natural frequencies and magnitudes of risk. This creates an excessive focus on the pre-emption of risk,

while leaving out handling the implications of uncertainty, that is when the pre-emption risk turns out not be successful. As my critical analysis highlights, current decision-support practices seem to neglect preparing people to cope with both the uncertainties in the decision-making process as well as the potential implications that might ensue. This neglect comes into being via narrow conceptions of patient autonomy. In the context of medical situations with substantial uncertainties and potential outcomes that have previously not been considered, such as in the case of cancer screening decision-making, I argue patient autonomy may be better seen as a collaborative knowledge practice (Kukla 2007, 28), in particular with regard to the construction of preferences and values (Epstein and Peters 2009, 195).

Thirdly, following the insights derived from the critical analysis, the chapter will discuss the choice of issues which will be operationalised in the practice-based design research and discuss the respective design requirements. Attending to the choice of issues lead to the development of two prototypes. The first one is an interactive Venn diagram which focuses on facilitating an understanding of false-positive and false-negative PSA test results, and by way of that the inherent uncertainty in making sense of such quantitative diagnostic test results. The second one is a set of reflection cards, that aim to support the preference construction process which is necessary when aiming to cope with the uncertainty in diagnostic test results.

Informational health practice

In Chapters 1 to 4, a critical review and analysis of relevant literatures has been conducted in order to elicit a multiperspectival framing of the problem. These include online health information search and digital culture studies, social and cultural studies of evidence-based medicine, theories of cognition and design research. In the practice-based research, I will work on the topic of decision-support from these perspectives and how they provide different facets to the problem. To this end, the insights gathered from the critical analysis operate as a set of connected problems and thus thinking resources which my practice-based research is attending to. In what follows, I will selectively set these problems up to inform the practice-based design research project. Thus, this research draws inspiration, both from the phenomena of actual health information practice but equally from theoretical reflections such a health information practice is constituted and affected by.

Chapter 1 starts with the observation of a strong interest by some patients to be informed and participate in medical decision-making as regards their health. One particular and significant way in which such interest manifests is in

the frequent use of health information sources online found via search engines. According to the Pew Research Center, searching for health information online had become the third most common activity among internet users aged 18 and above in the USA by 2010 (Zickuhr 2010, 3). For many, health information found online affects their medical decision-making in terms of how to treat an illness or ask doctors new questions (ibid., 4). Given the prevalence of online health information seeking via search engines, the first chapter looks at the ways in which health information seeking is supported via the design of online search interactivity and, thus, can be assessed.

To begin with, I argue that the nature of satisfying health information needs is often complex. On the side of the information seeking person this is frequently complicated by an absence of specific domain knowledge such as not knowing adequate key terms to search with, having a differentiated understanding of the veracity of the many resources and actors present online, or indeed the specific information goal itself. One of the ways in which people engaging in information seeking today materialises problematically is in the identified user behaviour of 'thrashing'. Thrashing occurs when users continuously search with mostly similar keywords, thus exhibiting an anchoring bias. As a consequence, typically search results will not change much and, thus, leave users unsatisfied (Morville and Callender 2010). Such behaviour thus is exemplary of an absence of knowledge domain comprehension and the capacity to query the internet in a range of different directions. A similar lack of knowledge domain comprehension materialises when it comes to evaluating the relevancy of search results presented in light of ones information seeking goal. Recent research on the use of and reliance on highly-ranked search results (Pan et al. 2007) has highlighted that search engine users tend to trust artificially higher-ranked search results more than those ranked lower on the search results page, despite the fact that their abstracts seemed less relevant to the task. Thus, accountability for web search results, in the absence of specific domain knowledge, seems mostly to rely on the status authoring role through those search engines (Rogers 2009).

Following such empirical insights, Broder (2002) suggests differentiating between three different types of search behaviour, which is a useful taxonomy for the purposes of the present discussion. Firstly, navigational search queries, where users want to find the URL for a specific website. Secondly, transactional search queries, such as checking flight prices, which can be performed on a number of different yet specific websites. Thirdly, informational search queries to find information that may be present on multiple websites and where the search goal may not always be clear at the beginning but only emerges through the search process itself. Typically, search engine users exhibit very little problems with both navigational and transactional search queries. This is because, upon being presented with such results, they can

easily identify a specific URL or a set of URL's which will serve they search goals well. The empirical account of search engagement described is most problematic for informational search queries within a knowledge domain an information seeking person has little experience and comprehension of, such as the health and medical knowledge domain when a person is confronted with a new diagnosis. Such an understanding serves as evaluative criteria for analysing the design of search engine interactivity and its adequacy and usefulness within the context of health information seeking by non-medical experts.

My analysis finds that the current design of online search engine interactivity is strongly oriented towards an information-retrieval (IR) use scenario, in which users are assumed to know both their information goal as well as relevant key terms to search with, and upon presentation of a set of potentially relevant search results have little difficulty in locating the desirable and useful results. Indeed, such IR orientation has been further elaborated by way of algorithmic filtering of search results (Feuz et al. 2011). Such an approach however, further obscures the health information seeking process as the information seeking person is unable to understand or participate in the filtering process and the grounds on which its relevancy has been decided for her. Following my analysis of the design of online health information search interactivity in chapter 1, the nature of health information seeking interactions is argued to be badly supported by universal search engines today. Alternatively, an exploratory model of interaction is proposed as productive guiding the design of health information interactivity, both online and offline. Such a model of interaction entails learning about a knowledge domain, its concepts and actors as part of the interaction process itself, from which the informational goal may evolve and clarify over the course of multiple sessions.

In chapters 2 to 4, I have looked more specifically at the conditions in which such desired participation in medical sense-making and decision-making is supported by the larger health information ecology, and in particular with regards to cancer screening decision-making. To do this, I have focused on the analysis of both, the nature, quality and limitations of evidence-based medical knowledge production as it relates to individual health information seeking and sense-making, as well as the specific ways in which health information interaction and decision-making is currently supported within the offline medical encounter such as GP appointments. The analysis finds that the ways in which health information is produced and presented as well as the way in which decision-making as a process is framed further substantiate the potential relevancy and value of the proposed exploratory interaction model for supporting shared decision-making in context of cancer screening. The key

insights and arguments from the critical analysis will be reiterated here.

Cancer risk information is generated following the evidence-based medicine epistemological model. Specifically, based on epidemiological methods such as randomised controlled trials, evidence of cancer risk as well as the effectivity of diagnostic tests and therapeutic interventions is generated based on samples aiming to be representative of reference populations. Apart from recently identified epistemological issues concerning the quality of such evidence, more fundamental questions exist regarding its meaningful applicability at the individual decision-maker level. While epidemiological methods calculate a statistical risk across a reference population, at the level of the individual decision making person, such statistical average risk information may be said to bear little predictive value. This is due to the fact, that pathology at the individual level is more akin to a binary logic, one develops cancer or not. The use of probabilistically derived cancer risk-factor information can at best improve the 'average health of a population'. In prostate cancer screening, this is operationalised by means of the epidemiologically determined probable link established via a linear gradient smoothly connecting normality (being clinically healthy) with risk factors (with or without symptoms) and pathology as defined by measureable quantitative changes and corresponding probabilities of PSA blood values.

Unfortunately however, this comes at the cost of framing healthy people as at risk of future pathology, due to methodical limitations that materialise in the form of false-positive and false-negative test results, and, thus, over- and underdiagnosis of cancer. An example of a false-positive diagnostic test result and thus overdiagnosis occurs when a patient receives the "diagnosis of a 'cancer' that would otherwise not go on to cause symptoms or death" and which may nevertheless be treated preventively (Welch and Black 2010), while false-negative test results and underdiagnosis refers to the opposite scenario.

As has been discussed in chapter 3, the style of presentation of evidence-based medicine cancer risk information creates further complexities. This specifically concerns the use of natural frequencies and visualisations of magnitudes of risk. To reiterate, natural frequencies refer to base rates, and, thus, aim at preventing people from committing errors due to the base-rate fallacy. Focusing on an understanding of risk in contemporary cancer screening decision-support comes at the cost of obscuring the fundamental uncertainties involved in such decision-making at the individual level. In turn, this generates the illusion of being able to control the future while being unable to prevent alternative courses of one's health and the subsequent need to account for one's previous action and decisions (Esposito 2007). This relates to the scope and framing of contemporary decision-support processes.

As my analysis has identified, there is an absence of supporting preference construction for preference sensitive care decision situations such as cancer screening. Preference sensitive medical care situations refer to healthcare situations, for which there is no single best medical practice to follow, while the implications of the medical practices available can affect peoples' personal lives very differently. Thus, as Mulley et al. argue (2012), patient preferences matter in such situations. Unfortunately, medical professionals do not appear to be very good at diagnosing patient preferences (ibid.) either.

The understanding of informational health practice that follows from this, is one that highlights an asymmetry between the information needs of people and patients and the ways in which health information can be accessed and interacted with and, thus, may facilitate participation in shared decision-making. Furthermore, current informational health practice may be said to lack addressing patient information needs in a quality (in the sense of the distinctive characteristics it is made of rather than degrees of excellence) that is attentive to the full scope of potential implications such medical decision-making may require patients to consider and factor into the decision-making process. Instead of accounting for the irreducible uncertainty at the individual level involved in screening decision-making based on probabilistically derived epidemiological risk information, the focus is put on representations of magnitudes of risk, which are largely unintelligible at that decision-making level. Instead, coping with potential implications inherent in such uncertainty calls for preference sensitive decision-making. This includes questions such as who to involve and account for and which implications to be able to cope with and in which contexts. As the critical analysis highlights, such an understanding is not reflected in contemporary design of decision-aids.

To that end, the understanding that follows from a critical analysis of such informational health practice is that contemporary decision-support for cancer screening is primarily focused on an attempt to pre-empt risk. However, it omits to simultaneously attend to the underlying inherent uncertainty of such risk information at the level of the deciding individual, and more importantly, how the unpredictable effects of this may need to be coped with. It is as if the contemporary understanding of preventive medicine with regards to screening is only concerned with the biological aspects of prevention, irrespective of the potential psychological and sociocultural implications this may have for people. Thus, the attempt to prevent and pre-empt biological malignancy in diagnoses of prostate cancer may simultaneously actively generate a whole host of new problems such as anxiety, incontinency and impotency, which are currently not considered as relevant to be prevently considered, factored in and attempted to be prepared for coping with their potential arrival. My

research aims to foreground these aspects as part of the decision-making process. The ways in which this will be operationalised will be discussed below.

Having consolidated the understanding of informational health practice that follows from the critical analysis of literatures, I hope to have been able to convey the ways in which these subject areas are interestingly and problematically interrelated and provide a meaningful set of facets to the problem of cancer screening decision-support. In the following sections, I will clarify my position regarding the role of probabilistically derived cancer risk information for this research. Furthermore, I will outline which aspects in the critical analysis have been selected to be operationalised in the practice-based research component of this thesis.

Position on probability theory

The literature review and critical analysis has identified and problematized a range of issues inherent to evidence-based medicine and current understandings and design approaches towards decision-support. The analysis in the preceding chapters has particularly identified both valuable and critical aspects of probability theory regarding its role in medical epistemology. The critique is specifically focused on its indiscriminate application to support individual decision-making generally speaking, and in particular as it relates to preference-sensitive care situations such as cancer screening. To this end, this critique does not negate the role of probability theory in medical epistemology per se, but the specifically identified limitations inform the focus for the design research phase of this research project. This design research thus aims to address those limitations by researching how to most productively complement decision-support. Instead of focusing on a smooth relaying of risk-information such as seems to be the aim of using natural frequencies (chapter three) in decision-aids, this design research aims to explore how to design for uncertainty, which may be said to be (a yet underarticulated but) inherent and uncontrollable component of probability theory.

In the sections below I will discuss two of the issues identified in the critical analysis that will be operationalised in the design research phase of the project. Thereby, I will also consolidate the main arguments why the issues have been selected as desirable aspects for the design research and how they relate to the critique of probability theory in the preceding chapters.

Operationalization of selective aspects from the critical analysis

For the practice-based part of the research project, a selection of two focus areas has been made. As part of the design research, a set of prototypes have then been developed to start exploring and addressing some of the problems identified in the literature review. The first prototype is an interactive Venn diagram, aiming to support the understanding of false-positive and false-negative PSA test results. The second prototype is a set of reflection cards aiming to support prospective reflection and preference construction process. The following section discusses the selection of the two issues and the respective design requirements following from the critical analysis.

The first issue the practice-based component of this design research is concerned with is the understanding of the limitations of using epidemiological risk information in decision-support at the individual level. One of the ways in which this materialises problematically in decision-support today, is as a lack of understanding false-positive and false-negative test results and the implications this may have for an individual patient. A brief summary of the main issues identified and discussed in the critical analysis chapters seems valuable at this stage.

To start with, there are significant issues with decision-making patients being able to make sense of statistical information (see chapter 1). Risk information visualisations often represent magnitudes of risk rather than the actual random distribution of risk at the individual level (see chapter 3). To use magnitudes of risk is a questionable design decision in particular within the context of shared-decision making. The applicability of such risk probabilities at the individual level is generally questionable due the notoriously low discriminatory predictive value of such risk-factor information (see chapter 2). Furthermore, in the specific case of prostate cancer, there is a very high level of false-positive test results with the PSA test (see chapter 2) and thus resulting in the requirement to undertake a biopsy in order to clarify the veracity of the PSA test result (see chapter 4). The logic of improving the average health of a population by means of screening, and, thus, the resulting numerous false-positive PSA test results comes at the cost of framing healthy people as at risk of pathology. This in turn, may implicate a number of side-effects such as infections, anxiety, incontinency and impotency resulting from the repeated biopsies and potential preventive measures conducted affecting all people undertaking screening, that is the healthy and the pathological ones.

In aiming to develop epistemic interactivity for an understanding of false-positive and false-negative test results, this design research aims to render

more intelligible and meaningful both the specific meaning of probabilistically derived risk information as well as its limited predictive ability at the individual level. In the best of cases, it is hoped that such an understanding would allow to prospectively imagine the potential courses of actions deemed necessary to clarify borderline PSA test results as well as the downstream implications this may have for the individual such as anxiety, infections as well as potential overdiagnosis and corresponding side-effects such as incontinency and impotency. To this end, the prototypes envisioned would complicate and complement what probability theory may make appear as rationally desirable with a prospective reflection of the personal cost such pre-emptive risk management may come at. Thus, the design intervention does not aim to devalue the informational value derived from probability theory, but rather support an understanding of the underlying uncertainty inherent to it, and thus, may render a need for means by which its fallout may personally and socially best be coped with and accounted for more desirable. Following from my critical analysis in the preceding chapters, this is a limitation of probability theory, which in a context of shared decision making needs to be addressed so as to adequately live up to the demands of it. This brings me to a second issue.

The second issue the practice-based component of this design research is concerned with is supporting the personal preference construction processes in preference sensitive medical care situations, such as cancer screening. Given the limitations discussed with regards to the meaningfulness of probabilistic risk information at the individual level in the context of preference-sensitive care as well as the absence of supporting preference construction in contemporary decision-aids, such a design research focus seems most urgent and promising. As discussed in the critical analysis, this is relevant because decision-support based on epidemiological risk factor information alone, does not help prepare for coping with other than predicted or desired outcomes. Nevertheless, the ways in which cancer screening and its implications affect individuals in their personal and social lives varies greatly. As a consequence, preferences and values matter when making decisions about cancer screening. The information and decision environment is complex and patient autonomy may better be conceptualised as a collaborative knowledge practice. The patient may better be conceptualised as an active inquirer, supported among others by experienced physicians, who may be said to have the opportunity to develop vast experience with the kinds of social and moral concerns that may surface within their medical field of expertise, and, thus, may offer some valuable advice and direction on possible dilemmas (Kukla 2007). In the following section, I will discuss the specific requirements on the design process which follow from the literature review for each of the two

issues selected.

Focus 1:

One of the core requirements is to support a differentiation between understanding risk and uncertainty. While the former can be calculated at the population level, its applicability and intelligibility at the individual remains fundamentally uncertain. The design process and resulting artefact thus, should support people to grasp uncertainty. This in turn also means finding a way to undo the reliance on magnitudes of risk as a relevant information cue for individual decision making. Following the extended mind thesis, such interactivity should be enacted by the patient/decision-making person rather than just received as information through reading in a brochure or hearing from a doctor. Being able to switch from an epidemiological perspective to the meaning of a test result from an individual perspective is necessary. The ways in which such a perspective switch occurs in the best of cases should enable raising meaningful questions about each corresponding state, and in turn, such interactivity may be meaningful and productive as epistemic interaction. That is, interacting with the prototype allows to explore, interrogate and start grasping uncertainty in prostate cancer screening decision-making.

Focus 2:

Complementing an understanding of the underlying inherent uncertainty in prostate cancer screening decision-making is a need for means by which implications thereof may personally and socially best be coped with and accounted for. Thus, the design process should facilitate considering the information concerning screening and its potential implications in the light of ones own personal and social life. This in turn should support identifying ones values and help constructing personal preferences. Deliberating personal values and preferences in such a novel situation may best be supported through the design of artefacts which help identifying relevant 'cognitive scaffolds'. This may be in terms of people to speak to as well as the meaningfulness of relevant language to be used and explored, but also the artefactual environment used to support such deliberation. As part of such deliberative processes, identifying normative values in ones social environment, one might feel influenced by, is equally important. Lastly, the design should operate on a time scale that factors in delicate conversations with ones wider social network and being able to reflect on thoughts and comments made.

Given these requirements onto the design process, the design strategy followed is the research through design method [RTD]. The method will be introduced and discussed in the following chapter in detail.

Chapter 6

Design project and method

Introduction

As discussed in chapter three, the perspective taken in the visualisation and communication of risk information seems largely oriented along epidemiological principles and public health management. This makes it less relevant or, at least, insufficient for supporting individual decision-making. The visualisations seem unable to support the specific issue of understanding risk from an individual decision-taker's perspective, particularly with respect to the probabilistic character of such information and the randomly distributed nature of such risk (and in turn the false-positive or false-negative results). In effect, existing risk and uncertainty visualisations give a false sense of risk distribution which, from an individual perspective, does not apply; as in the case of a single event probability, the actual outcome is (more like) a binary as one is either affected or not.

Besides the identification and analysis of the fundamental problems discussed in chapters one on search interaction models, chapter two on evidence-based medicine and chapter three on risk information communication and models of decision-making, various less directly related resources fundamentally inspired and informed the design process. These will be discussed in the following sections as well as where directly relevant for the design process.

Inspiration for the design process

In deliberating about the design process and the expectations I had concerning the specificity and capacity of the tools to be designed within the problem space described, various writers articulated well such potentials. Clark (2009) himself suggested putting EMT to use for transformative purposes rather than replicative ones. With this he refers to a variety of newly-developed technologies, such as telepresence, often aimed at replicating in detail our existing forms of bodily action and experience. In contrast, he highlights, for example, e-mail as a new technology providing a variety of new and complementary functionalities to make use of written language. This includes allowing people to interact informally, swiftly and asynchronously, while keeping a trace of the communication.

Transferred to the context of prostate cancer risk visualisations, I take Clark's suggestion to mean moving away from the reductionist ways such risk information is verbally and visually provided today (see my analysis in chapter three), making it unlikely to support a transformative understanding of the potential personal implications of uncertainty in the screening processes. In contrast, Clark (2003, 111) is hoping for EMT-inspired and informed design to be "expanding and reinventing our sense of body and action [...] and, thus, allow to expand the types of engagement we enter to." The potential I see in this stems from combining Mol and Law's suggestion to privilege action with EMT's epistemic action to achieve such an expanded sense of body and action.

David Spiegelhalter, a statistician and Professor for the Public Understanding of Risk at Cambridge University, summarised succinctly his expectation of risk visualisations. For him, "the greatest challenge is to make a visualisation that conveys its own contingency and limitations" (Spiegelhalter 2011, 1400). This is a particularly relevant and desirable characteristic of an information visualisation when considered in light of prostate cancer screening, as argued in chapter four. This is because such contingencies and limitations need to become part of the cognitive and social processes which are facilitated by the visualisation in the first place. As my discussion in chapter three highlighted, this is seldom the case in today's risk information visualisations.

The challenge of conveying the model's own contingencies and limitations is an aspect also taken up by Fuller in relation to decision-support systems more broadly. He specifically refers to the redistributive mechanisms concerning "the way it can help displace and redefine expertise, valid knowledge, the landscape of choice and the rationalisation one makes of it" (Fuller 2012, 134). A more detailed discussion of the relevance of these inspiring perspectives will follow as part of the prototype development process.

A slightly different view on the question of expertise, chance and how we account for decisions has been articulated by Bursztajn et al. (1981). For them, uncertainty plays a much larger role in medicine than has previously been acknowledged and, thus, needs to be accounted for accordingly in the medical methods. They proposed doing so by explicitly differentiating between medical situations

that are largely governed by chance and situations, in which skill can be decisive. By 'chance' situations we mean those where nature must take its course; by 'skill' situations we mean those where human action can influence the outcome. When we speak of 'skill' vs 'chance,' we are not talking about the degree of predictability that a

situation exhibits, but about the degree of control one may exert over it through the exercise of skill.

(ibid., 196)

This view of chance and skill is precisely applicable in the context of prostate cancer screening. While the result of a PSA test is to a large degree governed by chance, undertaking a biopsy skilfully can actually identify and diagnose a person's health situation.⁹³ Interestingly, the authors argue that a common error in medical decision-making is to regard chance situations as skill situations and vice versa. As is the case with prostate cancer screening, to see the PSA test as a skill situation and disregard the substantial level of unpredictability that plays into it comes down to a form of defensive medicine. The problem with such an attitude is obviously that it neglects to factor in its own performativity, potentially manifesting in iatrogenic effects discussed earlier. By explicitly and actively appreciating the role chance plays in some medical situations, Bursztjain et al. (ibid.) hope to enable both doctors and patients to use and refine their decision-making skills to live with anxiety instead of allowing anxiety to lessen their skills. For them, by actively acknowledging chance situations and 'gambling' on them, patients not only gain experience but also

a better sense of the probabilities, that is, which contingencies are controllable and which are not. They will then be able to make better choices in the future. Beyond this, the very act of making an informed choice can give a patient a sense of being in control, which in itself may reduce the stress of illness and with it the physical as well as emotional damage suffered by the patient.

(ibid., 203)

In the context of PSA testing and the question of undertaking biopsies, the 'gamble' refers to the personal decision on how often to test, how to interpret the results and, consequently, how often to undertake a biopsy. Should a person having had a negative biopsy after a 4.1 PSA test undertake another biopsy when the most recent PSA test result was 4.2 or 4.3, or wait until it has risen beyond 5.0?

⁹³ Even if still with some degree of uncertainty (personal communication with Felix Huber, 25. 02. 2014).

Methodological considerations and research through design method

Before discussing the actual design process, this section will set out and discuss the methodological approach taken for this research. Research through design (RTD) is a research method which privileges the exploration of a research question through the articulation and evaluation of designs. As such, the designer is more cautious in terms of approaching a design problem in ways that may prove productive in a person's future situated uses (Suchman 1987). As Gaver (2012, 940) argues, RTD is generative rather than comparative in nature, hence interested in the question of 'what might be' rather than designing 'the right thing' (Zimmermann et al. 2007, 498).

The method was first given its name by Frayling (1993) and has been further articulated by Buchanan (2001) but may also be said to be strongly influenced by Schön's *knowing-in-action* (1983). Schön, among other things an educator, developed his arguments for knowing-in-action based on what he termed *professional artistry*, by which he meant a competence in practice that does not depend on the ability to articulate such actions in conscious ways. Instead, such skilful competence is developed by learning "to appreciate, directly and without intermediate reasoning, the qualities of the materials that we apprehend through the tacit sensations of the tool in our hand" (ibid.). This relates to a form of knowledge in which meaning is deeply intertwined with our embodied interactions with the environment; in the context of a research project, RTD explores understandings in particular situations.⁹⁴

This perspective is closely related to the arguments made by Dourish (2004), Noe (2009) and Clark (1998, 2008) (as discussed in chapter four) and, thus, serves to highlight and further substantiate the relevance of RTD in light of the hypothesised productivity of EMT for exploratory interaction. This is because, within the context of this research, RTD may be said to be applying an EMT perspective onto the design research process itself. In this way, it is as if I explore the potential for epistemic action in a style of research practice that is in congruence with the very goal of identifying such epistemic action by way of designing it.

As Schön (1983) argues, knowing-in-action is guided by a kind of normativity of doing things right, be that riding a bicycle or catching a ball. In contrast to this, he refers to reflection-in-action as situations in which we notice that our usual action in a new context produces an outcome that does not meet our expectations. However, instead of reflecting on the outcome after the fact, in some situations we reflect "in the midst of action without interrupting it"

⁹⁴ Michael Polanyi, a polymath, referred to this as 'tacit knowledge' (Polanyi 1958).

(ibid., 54) and, thus, make immediate changes. What such reflection-in-action does is applying a tacit (Polanyi 1958) knowing-in-action combined with an inquisitive mind led, for example, by a specific research question.

The outcome of such a process, as Gaver (2011) argues, is generative in nature and, thus, the method is often chosen as a research method in contexts that are thought to exhibit a potential to create new knowledge. This happens, for example, through the annotation of the resulting designs (ibid.), which may take the form of prototypical artefacts or whole systems (Gaver 2012). For him “the resulting designs are seen as embodying designers’ judgments about valid ways to address the possibilities and problems implicit in such situations, and reflection on these results allows a range of topical, procedural, pragmatic and conceptual insights to be articulated” (ibid., 937).

Most recently the discussion within the field of Human-Computer Interaction (HCI) has turned to questions of developing agreed methodological standards of what constitutes “good design research” (Forlizzi et al. 2011; Gaver 2012). The goal of such a development would be “conceptual contributions to relevant and rigorous theory, which exhibits a level of extensibility and verifiability designers can apply in research in practice” (Gaver 2012; Zimmermann and Forlizzi 2008). Such calls for standardisation are symptomatic of more fundamental methodological discussions in the larger field of HCI concerning the introduction and legitimate acceptance of a third paradigm for framing interactivity (Harrison et al. 2007). While according to Harrison et al. (ibid.), the first paradigm (human factors or ergonomics) is largely oriented towards specific and opportunistic man-machine optimisations with little interest in theoretical generalisations, the second paradigm’s central credo is the information-processing model of mind. This allows for the efficient modelling, design and positivistic evaluation of man-machine interactions, and subsequently theorising.

The third paradigm, in contrast, considers interaction as “a form of meaning making in which the artefact and its context are mutually defining and subject to multiple interpretations” (ibid., 6; Sengers and Gaver 2007) and aims at supporting situated action (Suchman 1987). Approaching shared decision-making with an openness for situated action may create space for deliberation which may support the preference-construction process. As I discussed earlier, not all patients see themselves capable of meaningfully interacting with an informationally complex situation or may even fear being medically ill-treated when raising questions that may be perceived as challenges to the medical professional.

Within the context of this research, no relevant design work seems to have been produced from the perspective of EMT (Smart 2012; Smart et al.

2010). Yet, the generative nature of RTD promises to be most productive as its performativity is concerned less with questions of epistemology and more with ontological politics (Law and Urry 2005). The generative nature of RTD entails an interest not only in the creative reframing or actual generation of novel contexts, but also thereby debating and inviting reactions on what is ontologically relevant. In light of such a view on the role and potential productivity of RTD, “the assumption that shared standards are necessary, possible or desirable, are potentially repressive acts of ontological politics” (Gaver 2012, 943). This is because such standardisation would pre-emptively limit renegotiating and exploring ontological possibilities and, thus, politics in an unconstrained way.

Planned design process

Following the discussion of the various hypothesized problems and design potentials when introducing EMT in relation to prostate cancer screening in chapter four, the design process at first aimed at developing a way in which these varying issues could be integrated. This was thought possible by firstly creating the individual tools that address the specific decision problems, and, secondly, by developing an overarching narrative structure which would integrate the individual tools in a meaningful and coherent way. The hoped-for effect was that, through such narrative integration, the individual decision aspects could be synthesised and, thus, facilitate a consolidated decision concerning the underlying question of whether or not to screen.

The overarching narrative structure was envisioned as a decision process which would include screening as well as diagnosing phases in prostate cancer screening. The motivation to include the actual cancer diagnosis phase was thought relevant as it provides a feedback on the screening outcome. This feedback includes the problematic false-positive and false-negative diagnostic findings as well as the continuing difficulty in diagnosing malignancy of prostate cancer and, thus, making prognosis,⁹⁵ which may materialise in diagnosis such as *mild or medium-aggressive* cancer. The implication of this is the difficulty of deciding what treatment path to follow, if any.

An early mapping exercise of the many steps necessary to bring together the individual decisional components in developing an overarching narrative structure, as illustrated below, quickly indicated how extensive this

⁹⁵ Personal communication with Felix Huber, a GP, and Cédrix Poyet, a urologist, with whom I evaluated my decision support tools in February 2014.

would likely need to become.

Decision process - overview

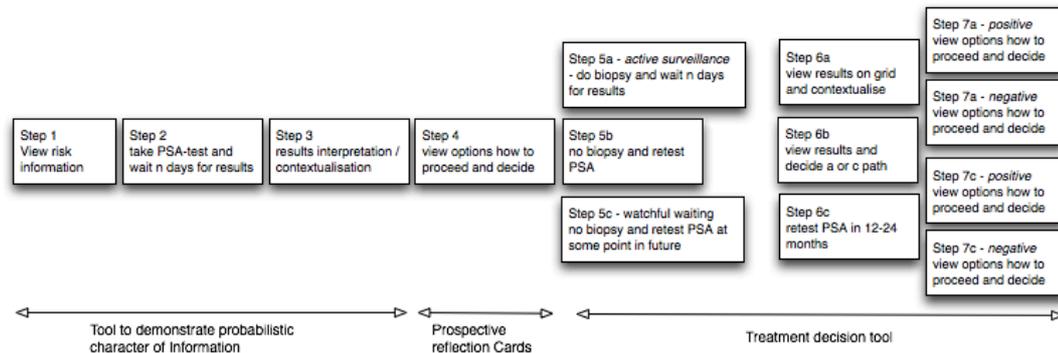


Figure 6.1: Prostate cancer screening decision process overview

In light of the existing constraints of doctor-patient consultations as well as the early stages of developing and experimenting with decision tools for epistemic action, such an approach seemed too laden with presumptions, leaving little space for appropriation by patients or doctors, hence an unlikely fruitful starting point for collaborating with medical professionals. Furthermore, when reconsidering the idea of an overarching narrative structure from the perspective of EMT, the idea seemed too invasive and not perceivably contributing to the goal of creating a relevant cognitive scaffolding open enough for others to participate in the design process.

From a RTD perspective, such an approach seemed to potentially change the kind of insights that can be gathered. Instead of researching the kinds of exploratory potentials that may emerge through the use of tools aimed at supporting epistemic action, a grand overarching narrative structure would potentially have generated more, or even too much, of a self-referential focus, that is too much of a focus on the tools themselves. Thus, it would more likely have asked whether or not the collection of tools through such narrative structure would better contribute to a consolidated decision. Clearly, this question is not at the forefront of this research, hence my decision to quickly abandon the approach of having an overarching narrative structure and, instead, focus on specific aspects only.

Due to the issues discussed, I reoriented the design process and chose to focus on two specific key aspects. Problem space #1 is the need to support an understanding of the probabilistic character of information underlying a PSA test and the implications of that from an individual decision-takers' perspective. Problem space #2 is developing tools for anticipating and reflecting on the ways in which the screening process might affect a decision takers' personal, social and professional everyday life, and how such preferences should be

factored into the screening decision-making process.

These two areas of focus were selected because they are thought to be core elements involved in the decision-making process. With regard to problem space #1, typically the epidemiological and, thus, probabilistic nature of the PSA threshold is often not well understood by people (Gigerenzer 2002, 4). This causes difficulties in grasping the meaning of an individual's PSA test result in relation to actual health and pathology and, consequently, understanding potentially false-positive and false-negative results and their implications.

Concerning the problem space #2, for a variety of reasons discussed earlier SDM is far from common today. Prominently among these reasons figures the absence of participation and involvement of concerned people and patients. A direct consequence of this of course is that, when resorting to a doctor's recommended path of action, people omit deliberating treatment options and the potential implications affecting their personal lives. Neglecting to deliberate these difficult aspects also prevents anticipating personal preferences and factoring them into the decision-making process. This issue has recently been labelled as 'preference diagnoses' (Mulley et al. 2012) and extends my earlier discussion concerning the identification and importance of preference-sensitive care (see chapter three).

These two areas of concern and interest were also selected as they are expected to provide a productive starting ground based on which this research can begin to understand how to design for and with EMT in mind. This is thought possible because the design challenges in both areas are very real and well understood by doctors, with whom I need to work closely in developing relevant tools. Thus, the expectation of what the designed prototypes should achieve in the best of cases was also very clear and will be discussed at the beginning of the next chapter. How this can be achieved, however, is at the very core of this research and design process and, thus, of the following documentation. Lastly, focusing on individual aspects separately was thought to allow for the more minimal and reserved approach that I felt is strongly desirable for the given context.

Problem space #1 - switching perspectives

As part of the exercise in developing an overarching narrative structure,⁹⁶ a

⁹⁶ As explained above, this was thought necessary to facilitate bringing the individual deliberations together to a consolidated decision on the question of whether to screen for prostate cancer or not.

number of ideas for tools addressing specific decision problems emerged. The following sections will discuss the insights gathered while iterating early prototypes of these tools with respect to their suitability for epistemic action. This early phase of the design process was instructive in developing a sense of which design aspects seemed key to focus on and refine as well as which were thought to be less suitable and productive for this research. This reflective design process allowed me to eventually hone in on two specific tool-ideas and, thus, make an informed decision against other potential prototypical tools. As will be discussed, this was a necessary step in being able to focus the design process and invest the necessary time and resources on the prioritised tool ideas.

Idea #1a – Stencil

The prime goal of the first idea to support an understanding of the probabilistic nature of PSA test results was to address the typically grouped style of visualisation of relative risk (see Figure 11 below). Such a style of visualisation fails to highlight the random distributive character of the risk of both pathology and, thus, of errors in identifying false-positive and false-negative results with overly sensitive markers, such as the PSA test.

As we can see in the visualization out of 1,000 men screened, the distribution of results is arranged in such a way so as to get a sense for the magnitude of relations between the different results. What this obfuscates, however, is precisely the random distributive character of the results. Thus, this style of visualization is unable to visualise the dynamic of the process in order to unravel which of the individuals with a positive PSA test result may in fact have a false-positive result which can only be detected upon conducting a biopsy. Some may argue that it is less important having a sense for the random distributive character of test outcomes compared to the relations of the diagnostic outcome and the number of men needed to be screened and treated. While the latter has a quality of a general micro-economic benefit vs. harm trade-off, the former hints at the *a priori* fictional (because probabilistic only) character of such a trade-off, which will necessarily manifest in a binary outcome one will have to deal with.

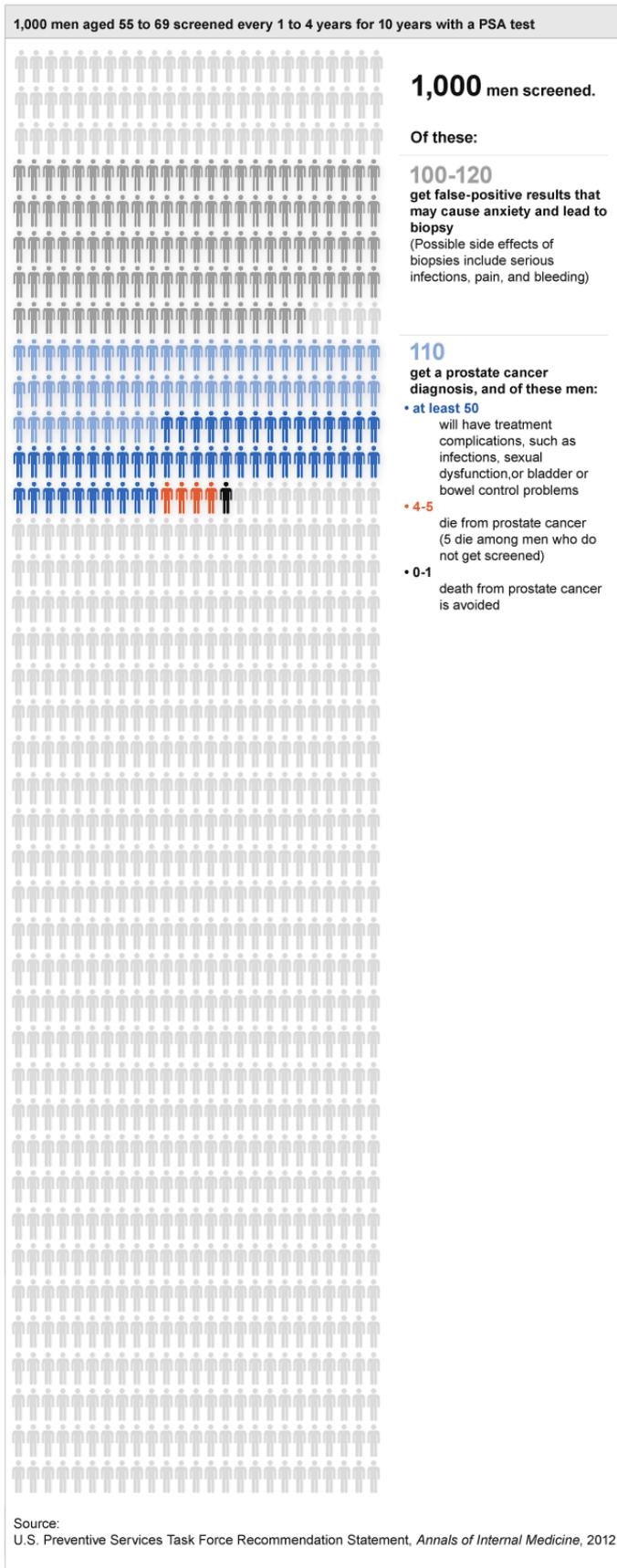


Figure 6.2: Grouped visualization style of prostate cancer screening outcome⁹⁷

⁹⁷ Image: U.S. Preventive Services Task Force Recommendation Statement (published on 17.

In order to convey a sense of the random distributive character of the PSA test outcome, the prototype was envisioned in the form of a stencil-based interaction. In a nutshell, a person would be given a paper with several rows of printed people-like icons on it. Each such icon would represent a person participating in the prostate screening process by taking the PSA test. The person would then pick one of those icons as being representative of himself and circle it. Furthermore, the person could circle other icons for prostate cancer screening relevant men in his social proximity and name them correspondingly. Next the doctor would overlay the stencil, which has holes on randomly distributed positions (but on the same row-and-icon-based underlying layout), so as to mark all those people who would find themselves with a positive PSA test result.

Importantly, these positive PSA test results would also include false-positive ones. In case the previously selected icon would be among the positive PSA test results, the doctor would then be able to contextualise the result by explaining that, while the result is positive, it could also be a false-positive one, and that the PSA test in itself is unable to differentiate among the two. In case the selected icon is not among the positive PSA test results, the doctor would be able to explain that this result could be either a true- or false-negative result. Furthermore, having other socially-proximate men relevant for prostate cancer screening circled and named was hoped to make the somewhat uncanny character of random distribution of pathology and error and its implications more broadly relevant.

Discussion of idea

A number of aspects of this idea seemed useful. Among them was its simplicity, as it would not take much effort to explain its application to doctors. In addition, it has similarity to the metaphor of throwing a coin when explaining the random distribution phenomena. Lastly, it was hoped the selection and marking process would make the quasi-binary⁹⁸ character of pathology at the level of an individual more clearly understandable. As discussed earlier, the use of quantitative risk information and its visualisation (marking only one half of a people-icon as pathological to correctly represent the statistics) in some cases led to misunderstandings; for example, people understood they would

07. 2012 in the *Annals of Internal Medicine*).

⁹⁸ See my discussion and problematisation on this in chapter two referring to Canguilhem (1994).

be experiencing symptoms or side-effects more or less often, depending of the quantitative risk information given, rather than being or not being affected by them (Gigerenzer 2002a, 4).

However, while potentially supporting a clearer understanding of the binary character of pathology, the form factor would not have allowed to more clearly visualise and enact an understanding of the simultaneous possibility and unknowability of a positive PSA test result to be either a true positive (pathologic) or a false positive (being healthy). Yet, this very aspect is at the core of what needs to be understood about the capacity and limitations of PSA testing. Thus, I felt that, since this aspect would not have been supported by the interaction with the tool and had to be explained by the doctor, the tool was limited in its potential for epistemic action beyond explaining the binary character of pathology.

Furthermore, the interactions with it were mostly to be performed by and necessarily but unproductively repetitive for the doctor, thus less suitable to support epistemic action for the actual decision taker. Also, its very form did not seem to allow to usefully include the issue with false-negative test results. Moreover, I was sceptical that the necessity of having a person initially select the icon representative for him would be understood as a credible simulation of the process. After all, the PSA test does operate on an actual biological reality rather than on an only simulated randomness; as a result, the perceived and speculated risk was that the simulation would be seen as an abstract mental exercise only, without much actual relevancy in assessing a person's susceptibility to developing prostate cancer or having it.

Nevertheless, iterating and prototyping the idea did allow for a number of insights that were productive for the continuing EMT-inspired design process. Firstly, designing for epistemic action by definition makes it necessary that the action is to be performed by the decision taker and not the doctor who is well aware of the underlying problems. Thus, the tool should be designed in ways that shift the interaction to the relevant person. Furthermore, while simplicity is useful, this should be employed with great care. In the above example, it may have been too simplistic and, thus, reductionist, as the design articulated more strongly the binary character of potential pathology and seemed less potent to facilitate an interactive understanding of true- and false-positive results. Additionally, the simulative character of this prototypical idea would have potentially diverted too much attention away from the epistemic action itself by requiring understanding and accepting its fictional set-up in the first place. The danger of this might have echoed in a lack of relevancy on the part of the decision taker in a very real context as well as a lack of acceptance on the part of the doctor.

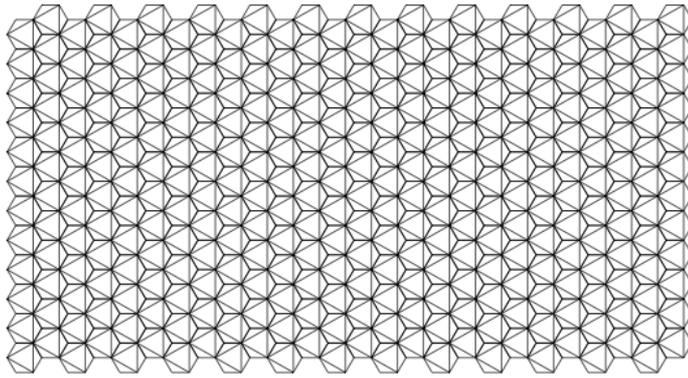
Instead, designing for epistemic action within the informational complexity of this context seems to require precisely orienting the interactional design to prioritised issues only, so as (again) not to overload the tool in the first place. Furthermore, by aiming for designed interactions to facilitate epistemic action, the specificity of such interactions should be anticipating and allowing for multiple interpretations. That is to say, the design of epistemic actions should not be over-engineered towards one specific way of conveying a specific aspect but, rather, aim at allowing for a degree of underspecification in its design.

As a working hypothesis, I believe such a design approach might be better suited to evoke meaningful and epistemic interactions. This is because, when cognition is supported through the environment, it is likely that people are used to a variety of scaffoldings that they may have used in the past; consequently, there is no one solution for all. Lastly, by leaving open an interpretative space, such a tool would potentially more usefully elicit a doctor's perceived role in contextualising and responding to possible interpretations.

Idea #1b - Grid

In reflecting on the limitations discovered in the stencil idea prototyping process, the aspect of productively designing epistemic interactions around the potential multiple meanings of a PSA test result was identified as a key point of focus and refinement. This specific aspect is at the core of understanding the uncertainties involved in the screening process and, thus, needed to be addressed.

Building on the previous idea, I experimented further with visual forms that would allow expressing such uncertainty. One such idea was based on using a grid as the underlying visual structure (see Figure 12). Again a person would have to select his position (see Figure 12 below) on the grid and, then, the doctor would reveal a randomly distributed set of PSA test results. What is different, however, is that the underlying visual form allowed for a layering of specificity of the results. The first layer would be the level of specificity revealed by a PSA test (see Figure 13 below). To uncover the second layer of specificity, the participant would need to be willing to undertake a biopsy (see image step 3 below). Only then would he be able to see which of the positive PSA test results were true positives and, thus, might require cancer treatment, and which ones were false positives and, thus, require no treatment.



Select your position

Figure 6.3: Idea #1b - Step 1: select position

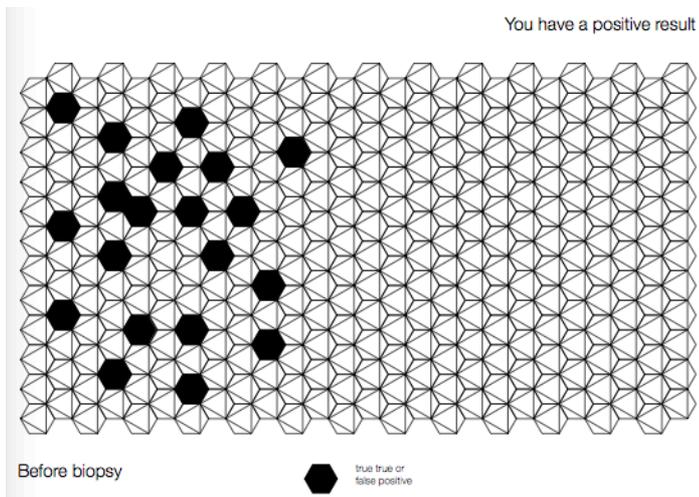


Figure 6.4: Idea #1b - Step 2: positive PSA test result

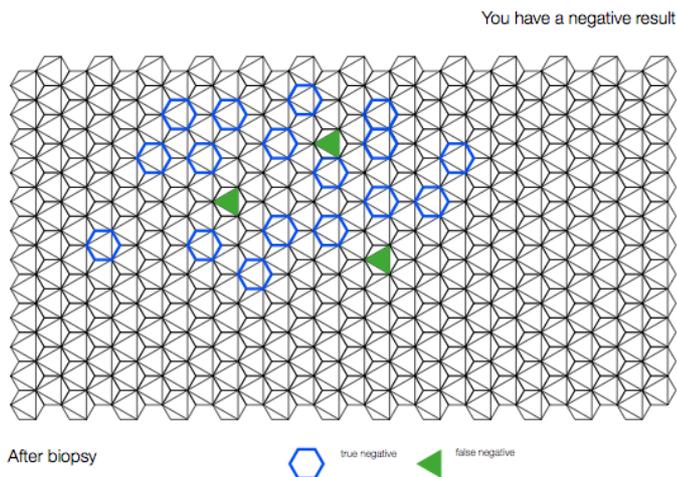


Figure 6.5: Idea #1b - Step 3: negative biopsy result

Step 2 would reveal which people (positions on the grid) would receive a positive PSA test result. By displaying a fully black hexagon, the PSA test

result would leave uncertain whether or not ‘underneath’ the black hexagon a true-positive (green triangle) or false-positive (blue empty hexagon) result would be found. This level of certainty would only become knowable if the person were to undertake a biopsy.

Discussion of idea

Compared to the stencil idea, this variant seemed to better encapsulate and visually articulate the uncertainty issues involved in PSA test results. While the visual form of hexagons did structurally facilitate to illustrate the duality of possible underlying health states of a PSA test result, its form seemed highly abstract for the context. Whether or not such a form would allow for appreciating the uncertainty issue involved seemed questionable. The reasons for such doubt were, once again, to be found in the limitations that the stencil idea exhibited as discussed above.

Additionally, the explanatory complexity in simulating the screening process and the meaning of each stage was increased, without a corresponding increase in the expected benefits through productive epistemic interactivity on the part of the decision taker. Furthermore, such an approach would most likely need to be developed digitally, and would require a narrative structure for it to make sense. This exercise reaffirmed the need to continue focusing on and experimenting with the visual form of uncertainty as well as finding ways to address the various identified issues.

Learning from the ideation and prototyping exercises thus far, it became increasingly clear that the tool for epistemic action I was aiming for should exhibit a number of key characteristics. Firstly, it should follow a minimal and underspecified design approach; with this I refer to the seeming complexity introduced with the hexagons which could assume multiple states. Improving the design in this respect would likely keep the requirements for explanatory embedding to a minimum and the tool itself lightweight and sharply focused on the core aspect of visualising and understanding uncertainty. Secondly, this specific thematic focus must be at the core of its interactivity. Such interactivity would most likely be epistemically productive if not just repetitive but instructive and transformative in understanding such aspects as the duality of health states of a PSA test result. Thirdly, such epistemic action should chiefly be enactable through a decision taker rather than having to be facilitated only by the medical professional.

Idea #2a - Venn diagram (mapping)

This emerging set of design criteria led me to experiment with a visual mapping of health states already used in chapter two to initially clarify my understanding of the historic views on the relations of health and pathology. Essentially, what I did there and continued to do here (see below) is applying EMT to the design process by thinking through visual diagrams. This was done by creating them and, thus, clarifying the meaning of the relations in them, or, as Schön (1983) would call it, performing reflection-in-action.

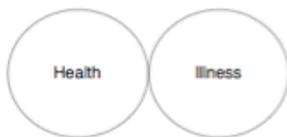
Pre-Broussais phase

During this phase it was held that physiology and pathology were two distinct domains of phenomena and knowledge. Consequently, it was thought that having knowledge about one of them cannot inform knowledge of the other. Signs and symptoms were indicators of pathology.



Broussais phase

Broussais conceptualised the normal and pathological as different states on a gradient of the same kind. The difference between the two states, according to Broussais, can be explained by changes of levels of intensity. Pathology, thus, was seen as an intensification of normal states.



Risk factor phase

While under Broussais' conceptualisation the pathological link is direct and causal, under the risk factor paradigm it is probabilistic and its aetiology only partially

understood. Pre-pathological states are symptomless and exhibit mild levels of risk factors. In pathological states, quantitatively measurable risk factors are always probabilistically pathological. The logic of health seems to morally implicate a reduction of risk factors.



Figure 6.6: Historical view of relations of health and illness

The reflexive use of EMT may be seen as an attempt to understand and critically reflect on the meaning of one's performed tacit knowledge practice. Or, in other words, as an active exploration of the possibly multiple meanings of one's performed action. Such a visual and epistemic mapping exercise was helpful and clarified the underlying uncertainty issue (see also discussion in chapter two). What was needed now was finding ways to build on this by making it understandable by interacting with it.

A way to engage a decision maker in such interaction while keeping the design minimal and focused was by having him map the result of a PSA test onto such a Venn diagram. For this I envisioned the simple exercise whereby the doctor would use such a Venn diagram and instruct the person to use pins and map the PSA test result onto the Venn diagram. To do this, the person would be given coloured pins for positive, negative and false-positive test results, the numbers of which would correspond to the epidemiologically known distribution of PSA test results. The doctor would then be able to observe where the pins were mapped and, if necessary, facilitate through discussion what the overlapping zone between health and pathology means. By having a person map his understanding of these states, this approach would serve as a simple test of their understanding, allowing thus the doctor to intervene should there be misunderstandings or questions about it.

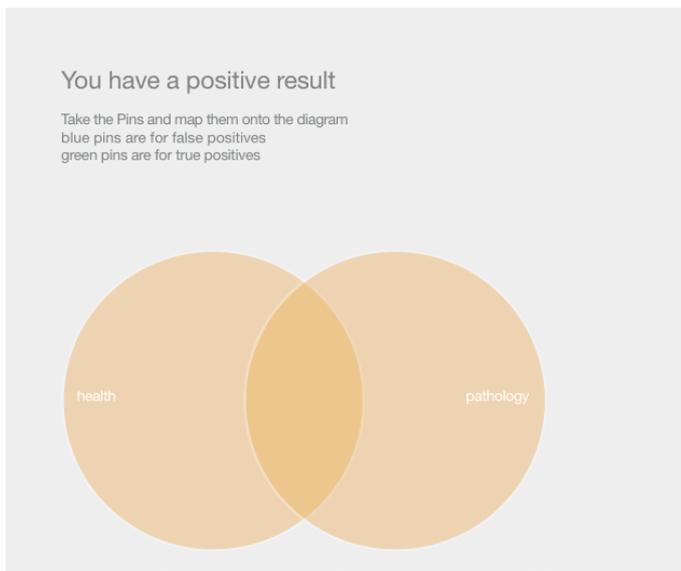


Figure 6.7: Idea #2a - mapping of PSA test result

A second and more refined version of the diagram included a scale for the PSA test result. This version would have allowed mapping an individual PSA test result, such as 4.3, onto the diagram. By way of this quantitative specificity, I envisioned the discussion between doctor and decision taker would be even more concrete concerning the localisation and meaning of false-positive results. Instead, the tool would allow to precisely locate a real or imagined PSA test result, such as 3.9 or 4.3, and, consequently, discuss the seeming pathophysiological urgency such borderline results seem to suggest (Greene 2007, 112), given that 4.0 is a widely used threshold indicating potential pathology.

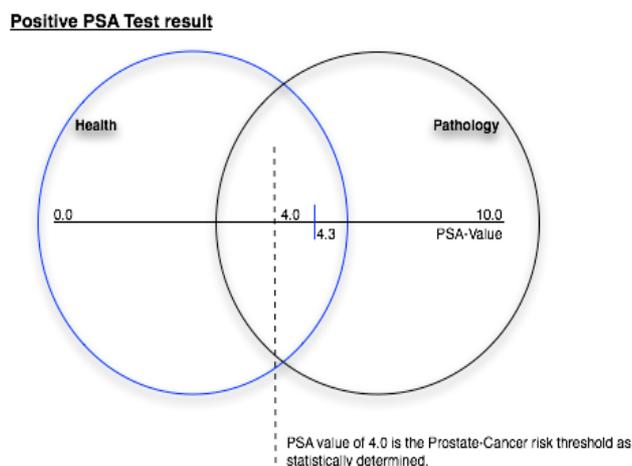


Figure 6.8: Idea #2a - mapping of PSA test result - epidemiological perspective

In order to clarify and problematise the meaning of such borderline results and the potentially felt uncertainty whether or not to undertake a biopsy, at this very point the tool should allow for the mentioned transition from the epidemiological perspective towards that of an individual person. This was delivered by another visual Venn diagram that illustrated such a false-positive test result by moving the relevant health and pathology circles:

False-Positive PSA Test result

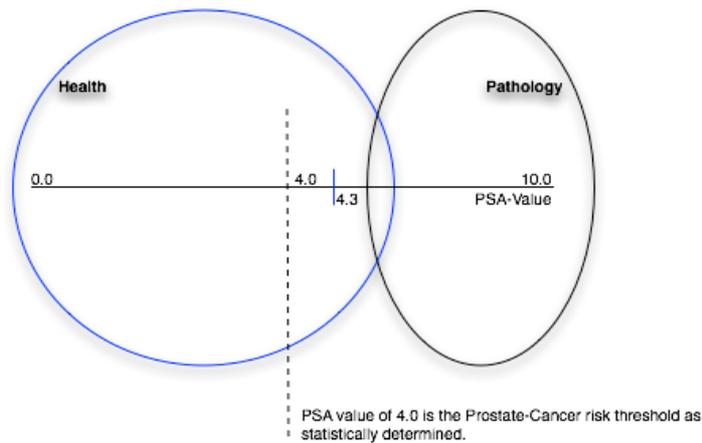


Figure 6.9: Idea #2a - mapping of PSA test result - individual perspective

The insight I crystallised from this exercise was that the Venn diagram should be interactable. Thus, the relation between the scale and the circles in the diagram needs to be changeable. I experimented with this requirement in the form of a paper prototype first.

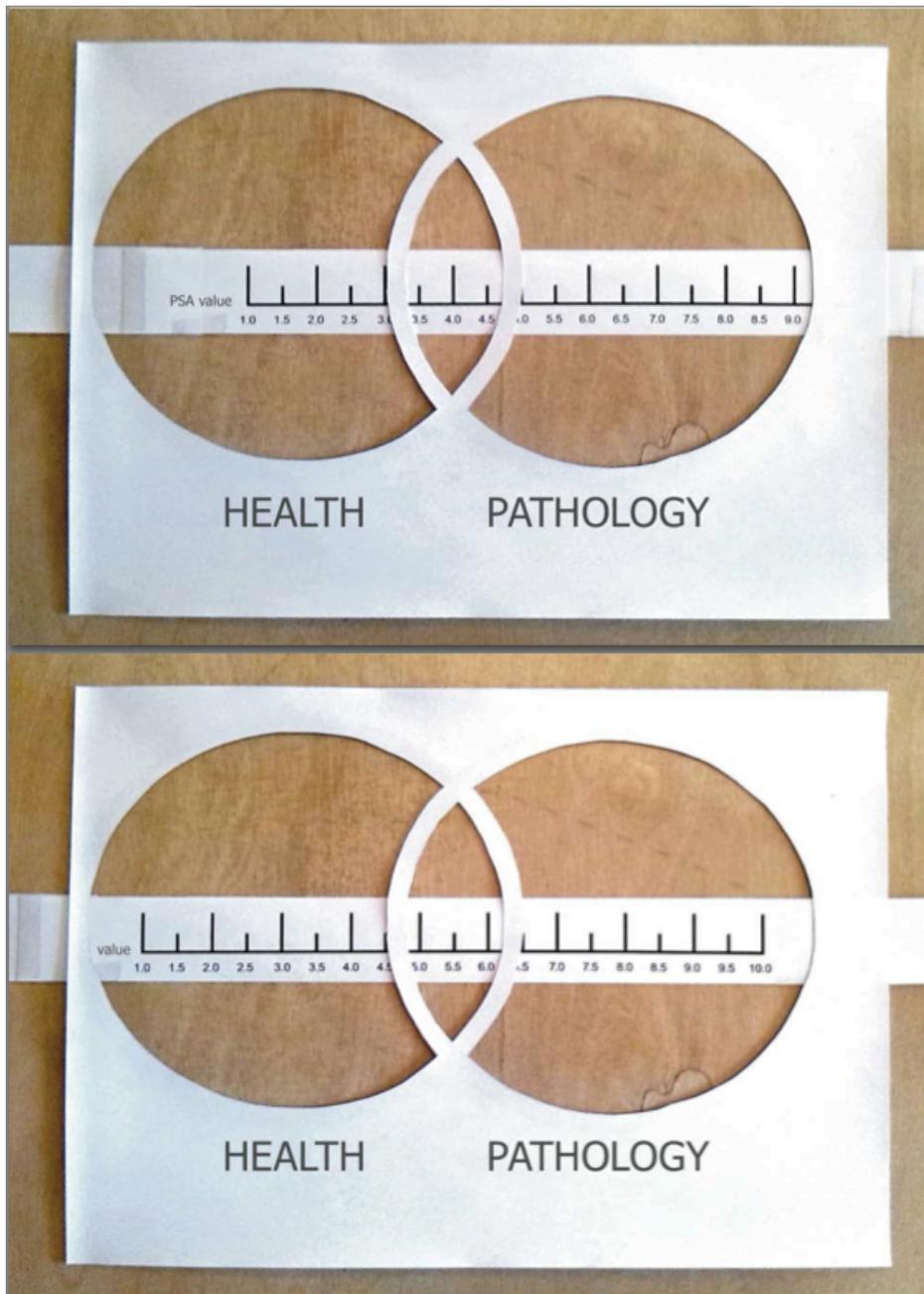


Figure 6.10: Idea #2a - mapping of PSA test result - movable scale

The paper prototype was designed by fixing the position of the circles and having the scale movable. This allowed to reposition the boundaries of health and pathology as, for example, in the case when a person had a 4.5 PSA test result and, upon conducting a biopsy, had learnt that he did not have prostate cancer as the PSA test result was a false-positive one. However, the fixed positions seemed problematic and not supportive in fully explaining such a potential situation. This is because, in contextualising such a negative biopsy result, the health and pathology circles for an individual could theoretically be anywhere from 4.5 to 10.0 and, importantly, from the perspective of an individual, no longer overlap after the biopsy. Emerging directly out of this was

the question of how large the problematic and, thus, overlapping zone between health and pathology should be.

I addressed this difficulty in a preliminary meeting with one of the medical professionals I planned to evaluate the final tools with. From a methodological perspective, the meeting was aimed at generating an interest in my work and the subsequent expert evaluation of the decision-support tools. Furthermore, the goal of the meeting was to ensure that my design work was oriented along relevant problems from the perspective of clinical practice as well as clarifying technical questions, such as the negative biopsy range above.

To my surprise, my conversation with urologist Cédric Poyet from the Prostate Carcinoma Centre at the University Hospital quickly revealed that, for him, the problematic zone for false-positive PSA test results was in fact very large and ranged from almost zero to 10 on the PSA scale.⁹⁹ This assessment was also confirmed in the evaluation with Felix Huber, GP and president of a network of group practices counting 250 GP's. Both of them also confirmed that this was an important issue with patients in their medical practice concerning prostate cancer screening. From an epidemiological perspective, this is little surprising, as it simply stems from the statistical variation which is the source of defining the mean and, thus, a threshold as a cut-off distinguishing between health and (potential) pathology. This is precisely the limitation of using such epidemiological data for individual decision-making, as it suggests a calculable likelihood of pathology inferred from a population as applicable to an individual's risk as meaningful information for decision-making (Politi et al. 2007, 682).

Discussion

The advantages of the mapping approach were that the interaction was shifted to the decision-making person, allowing, thus, for the doctor to potentially take on a more observing and facilitating role. Moreover, the design approach is utterly simple and requires very little explanation of how to use it. However, when using it myself and reflecting on it, I still found myself dissatisfied with the epistemic potential of the interactivity. Essentially, it seemed that, instead of generating an understanding by way of interacting with the tool, this interaction design approach required mapping an already existing understanding a person has onto a visual diagram. Thus, the act of interacting with it was again largely repetitive and would at best provoke the question about the

⁹⁹ It is important to note that this is specific to the PSA test.

overlapping zone of health and pathology states without necessarily clarifying it much through the interaction itself. This, I believe, was due to the static form of the diagram inherent in both the mapping exercise and the movable-scale paper prototype, which only had the capacity to represent the epidemiological perspective and not the actual quasi-binary state of health or pathology of an individual at a specific point in time (such as when making a biopsy).

The overlapping zone of the Venn diagram silently represents both the implications from overly sensitive markers, such as a PSA test, as well as a specific epidemiological view. The latter requires and assumes a long-term perspective rather than that of a specific point in time by linking current biological states to future probable pathology exactly by visualising these states as overlapping in the Venn diagram. What I thought was needed on top of the advantages of this approach was finding a way in which the interaction would allow a shift between an epidemiological and individual perspective in relation to a PSA test result.

My view on the necessary next step in developing the prototype was confirmed in a meeting with Cédric Poyet where I presented the Venn diagrams and discussed them with him. He explained that, in his daily consultation practice, he often thought about creating a tool which would allow him and his screening participants to track multiple PSA values and biopsy outcomes over time. This would not only allow them to see progression over time but also better understand a participant's likely individual relation of health and pathology beyond the epidemiological threshold suggested.

Idea #2b - Venn diagram (interaction)

The Venn diagram mapping approach fulfilled some of the desirable criteria, such as following a minimal and underspecified design approach, hence keeping informational overhead to a minimum. Furthermore, this approach also shifted the interaction to the relevant person, i.e. the decision taker. The necessary next step was to improve the design of the actual interactivity and have the thematic focus of understanding uncertainty through interacting with the tool at its core. While this was partly achieved with the movable-scale paper prototype, this approach was limited in its ability to illustrate the full spectrum of the epidemiological and individual perspectives. To improve the potential for epistemic action through interactivity, the next prototype needed to be able to fully address this.

Having realised from the paper prototype that a movable scale is insufficient to address this, the next prototype experimented with the

manoeuvrability of the health and pathology circles of the Venn diagram, while keeping the scale fixed. This approach, it was expected, would finally allow for the necessary flexibility to visualise both the epidemiological (overlapping circles) and the individual person's (non-overlapping circles) relations between health and pathology, and in this way support contextualisation of the meaning of such relations by the doctor. In designing this prototype, a number of elements were thought necessary to incorporate.

First of all, the ease of use for changing the relations between health and pathology is paramount. The design should allow for fluidly altering these relations, so as to be able to quickly switch between the different views. Furthermore, the form factor should anticipate and allow for note-making, so as to use the tool over a longer time and possibly multiple PSA test results. With regard to such a longer time-frame for its context of use, the idea for the tool was to not only support the discussion during the consultation but, also, to be seen by doctors as something they can hand over to their patient for home use once they have explained how to use it. Thus, space for instructions on how to use it should be allocated accordingly.

Given these requirements, and keeping in mind that the age-range of the prostate screening population is between 45-70 years, the overall size of the tool should enable both portability and easy handling and reading of its contents,. Lastly, with a view to winning over doctors for an evaluation with actual patients, the tool should be seen as useful and professionally produced for a larger scale evaluation as well as designed in a way that it can be economically produced and quickly made available.

In order to meet these increasingly ambitious requirements while satisfying the criteria for the tool to be seen as a minimalistic working prototype, I worked together with the Japanese product designer Kione Kochi. This collaboration was methodologically significant because Kione had the product design skills to experiment with various ways in which the imagined interaction could be physically built. This, in turn, allowed for discussing the benefits and drawbacks of more detailed aspects of the design. Furthermore, she could design the tool so that it would be robust and professionally producible in larger numbers, should a patient evaluation follow.¹⁰⁰

In setting up the design of the final prototype, I briefed Kione on the prostate cancer screening context, the core problems involved in PSA testing, and the need for undertaking biopsies to clarify the uncertainty stemming from PSA testing. Furthermore, we discussed the Venn diagram with the different states and relations of health and pathology and their meaning. The paper

¹⁰⁰ After the completion of the thesis and in subsequent collaboration with the medical professionals.

prototype with the movable scale was demonstrated and its limitations explained. Lastly, I ran through the requirements discussed in the previous paragraph and the criterion for the form factor to retain an appearance of minimalism and underspecification.

The design process ran through three stages: firstly, a draft conceptual design; secondly, an implementational clarification of the detailed interactions possible; thirdly, a review of the working prototype. Each of these will be discussed in the following sections.

Draft conceptual design

To start with, the basic interactivity was conceptualised as two sliders which would allow to easily manoeuvre the position of health and pathology states and, thus, their relation from overlapping to adjacent. This was envisioned to work as two translucent layers, with printed and positioned elements positioned on them accordingly. The scale of PSA test results was set to a fixed position in this prototype. Given the design brief, these layers would be integrated in a form factor of about postcard size and made of cardboard. The foldable cardboard would be designed in such ways so as to accommodate space for instructions and notes. The conceptual sketch below visualises the design we aspired towards.

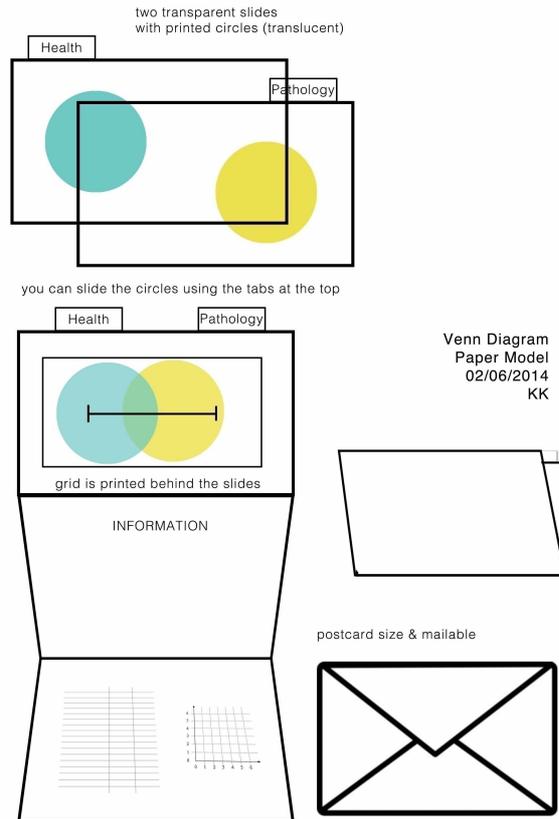


Figure 6.11: Conceptual design for interactive Venn diagram

Prototype production

When producing the cardboard prototype, a series of detailed questions surfaced that required implementational clarification of the detailed interactions possible. This specifically concerned clarifying the definition of the minimum and maximum zones for health and pathology states as well as the necessary and possible overlapping zone for these states. From reviewing the relevant literature, these values seemed neither easily nor non-ambiguously definable. Based on the literature and as a working hypothesis, for the epidemiologic view I assumed the minimum healthy-only zone to be from 0 to 2, the minimum pathological-only zone to be from 8 to 10, and thus the overlapping range to be from 2-8. Clearly, this was an aspect that needed to be clarified with the medical experts before such a tool could be evaluated with actual decision-making persons.

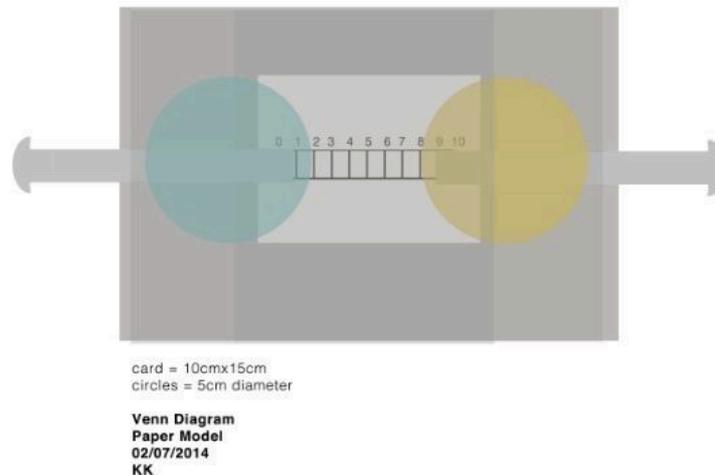


Figure 6.12: Interactive Venn diagram - minimum healthy and pathological zones

Review of the working prototype

Once the working prototype was available, I conducted a review of it. One of the aims of the review was to examine whether the requirements and criteria that evolved throughout the design project could be integrated in sensible ways. From my perspective, this approach worked well as the working prototype was both very minimal and its interactivity fully focused on the core issue it aimed to address, i.e. the shift from an epidemiological to an individual person's perspective on the relations of health and pathology.

From an EMT perspective, the way the prototype was to achieve this was by its ability to change the resolution of the mapping from the multiple possible states of a person as per the epidemiological perspective of a PSA test result (overlapping circles) to the perspective of an individual after having conducted a biopsy (non-overlapping circles). For example, upon receiving the PSA test result, the doctor would explain the test result in terms of overlapping health and pathology circles, indicating the possibilities for false-positive and false-negative results. In order to resolve such uncertainty, a doctor can further explain that a biopsy would need to be conducted. The result of the biopsy would be to switch from the epidemiologic (overlapping circles) view to the individual view and this would happen by moving the slides to be adjacent and, thus, the actual binary meaning of an individual's test result to be displayed. This epistemic move is the diagnostic equivalent of undertaking a biopsy as explained earlier.

Due to the minimalistic design approach, there seemed no need for

additional explanatory information as was the case with earlier design ideas and prototypes. Furthermore, the tool with its form factor allowed for easy handling and use over a longer time-frame, hence marking multiple PSA and biopsy results. Lastly, the tool could be designed in such a way that it would seem professionally produced. This led me to the decision that this should be the final working prototype to evaluate with relevant medical professionals in the field of prostate cancer screening.

Moreover, in reviewing the working prototype, I prospectively reflected on the planned evaluation with doctors and what I expected would happen in terms of their reaction to it. In other words, I tried to reflect on whether the tool would be well understood and appear useful in supporting their constrained patient consultation situation when explaining prostate cancer screening and its possible implications. At the heart of this was of course the question of whether the tool would allow to visually, interactively and meaningfully illustrate the issue of false-positive results.

A concern I had was that the use and conceptual understanding of a Venn diagram may not be broadly known and, thus, may not easily facilitate epistemic action for understanding false-positive test results. Generally, I wondered which aspects of the tool would seem most useful, problematic, limited or in need of further attention and modification. Among these was the question of whether doctors would see the tool as something they could hand over to patients once they had explained how to use it. This would be key for it to become a useful tool for people themselves in making sense of test result information should they decide to start the screening process.

More peripherally, and partly more as an aesthetic question, I wondered whether the texture of the health and pathology zones could be made more meaningful. This could potentially be achieved by exchanging an abstract pattern with the application of a man-like symbol instead. Given the early stages in designing for epistemic action and deliberately following a minimalistic and underspecified design approach, it seemed prudent to expect these questions to become partly answered when evaluating the tool with relevant medical professionals.

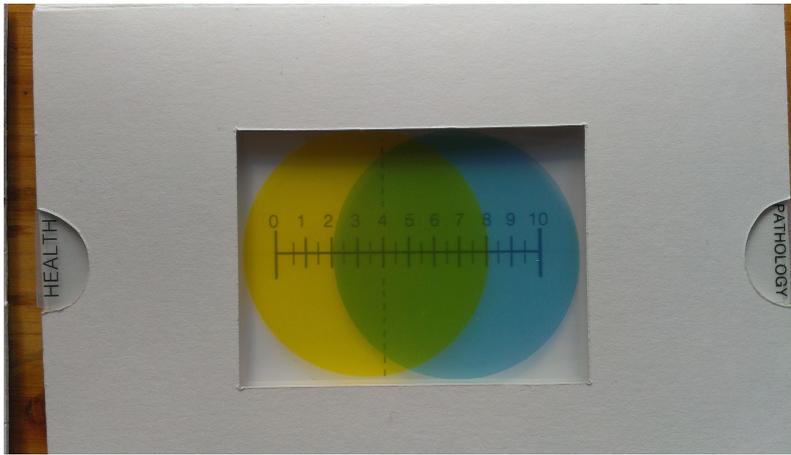


Figure 6.13: Interactive Venn diagram final prototype - epidemiological perspective¹⁰¹

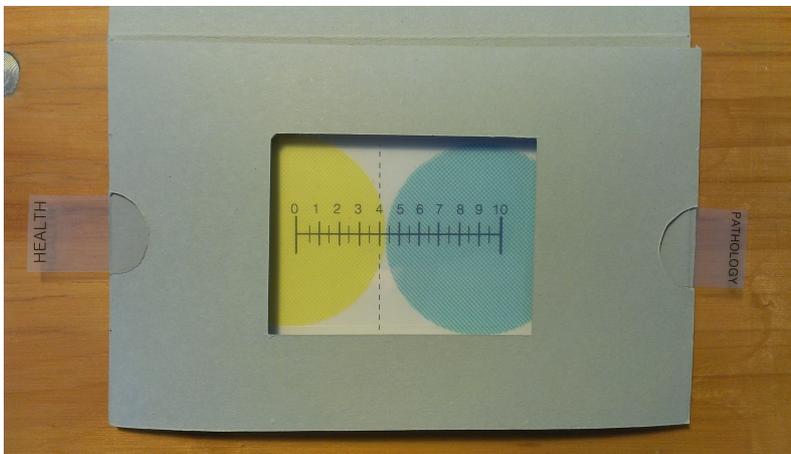


Figure 6.14: Interactive Venn diagram final prototype - individual perspective¹⁰²

Before such evaluation could begin, I envisioned a second tool to facilitate the screening decision-making process. This second tool, as discussed earlier, was aiming at addressing the issue with people's tendency not to participate in shared decision making and, thus, neglecting to deliberate personal preferences and factor them into the decision-making process. The following section will discuss the relevant design process focusing on this aspect.

¹⁰¹ The prototype is the size of a C6 Envelope, made of sturdy cardboard on the outside and with coloured but transparent slides on the inside. See also Appendix 2.

¹⁰² The levers on either side allow for the shifting of the overlapping zones. The left lever is for 'Health' and the right lever is for 'Pathology.'

Problem area #2 - preference diagnosis in shared decision making

As mentioned in the introductory section of this chapter, the second problem area I decided to focus on stemmed from patients' non-participation in shared decision making and their tendency to defer this to doctors and follow their suggested path of action.¹⁰³ Such behaviour is problematic as it omits deliberating on treatment options, the implications these might have for one's personal and social life, and the potential for retrospective regret after having followed a certain treatment path and only then learning about alternative possibilities (Mulley et al. 2012).

With regard to the specific context of prostate cancer screening, some of the potential iatrogenic implications of the screening process may manifest in psychological, physical and social realities. At the level of PSA testing, the uncertainty of having a false-positive or false-negative test result may make some people anxious (U.S. Preventive Services Task Force 2012). Conducting a biopsy, i.e. a surgical intervention usually undertaken in a hospital or clinic, additionally carries the risk of bacterial infections, which is increasingly problematic in hospitals and, in some cases, even lethal (Gale 2011). In the case of deciding to treat early cancer, potential side-effects of such treatment typically include incontinence and impotency. Such worries are reinforced by the fact that prostate cancer is among the cancers with a significant degree of overdiagnosis and overtreatment (Essermann and Thompson 2010; Welch and Black 2010).

I hypothesised that this contextual backdrop could productively be dealt with by prospectively reflecting on the potential positive and negative implications of the prostate cancer screening. I was, thus, motivated to develop tools for anticipating and reflecting on the ways the screening process might affect a decision taker's personal, social and professional everyday life as well as how corresponding preferences might be relevant for factoring into the screening decision-taking process.

Inspiration for creating a way to facilitate prospective reflection while building on the ideas of EMT was specifically found in the work of Elena Esposito, an Italian sociologist. In a 2007 essay, she compared the performativity and productivity of probability theory to realistic narrative fiction (Esposito 2007). To start with, she raises the question of how to deal with situations of uncertainty. She suggests that, in cases where we are inclined towards social consensus of decisions to be taken, a tool is needed which

¹⁰³ Personal communication with Felix Huber and Cédric Poyet, 25. 02. 2014.

would allow accounting for one's decision so they are also intelligible to other actors. For Esposito, this is precisely what probability theory provides, an instrument which promises future events to be calculable and, thus, possible to plan. This, in turn, would help us prevent regrets in the future, while maintaining consensus with other actors in the present.

However, Esposito argues that such planning cannot operate as a norm on the basis of which one can refuse to learn from actual experience. This is particularly true for those affected by the negative and considered less probable implications of such decisions (ibid.). For her, probability theory creates the illusion of *present futures*, future situations imagined from a person's present understanding, which remain unintelligible beyond the hoped-for and more probable outcome. Realistic fiction, however, allows to reflect on *future presents*, future situations informed by a broader understanding of possible outcomes, personal implications and what we make of them. Precisely as an instrument, its fictional character is utterly transparent yet realistic and, thus, actionable in its narrative content.

The extended mind thesis was thought potent in supporting such prospective reflection, among other reasons by way of building on its (hypothesised) role in our use of written and spoken language. In a nutshell, Clark (1998) argues that the act of speaking or writing helps us memorise and remember newly acquired information. Once spoken or written down, it becomes available for evaluation and self-reflection. Such evaluation and reflection processes help subsequently identify flaws in a plan we made, or critically reflect on the unreliability of our own initial judgment (ibid., 208).

Furthermore, and when anticipating longer interactive time-frames, Clark compares the potential of written language to the function that the aerial roots of mangroves have. These roots act over time by collecting debris (notes, questions, thoughts) and, thus, generate the fertile ground on which the mangrove tree can grow (understanding, insight, opinion). Similarly, through the use of spoken and written language by making notes, reflecting on them over time and being able to involve others may provide a useful ground for making an informed decision about complex issues, such as cancer screening.

Phase 1 - Prospective reflection cards

The way I approached these challenges was by creating a set of prospective reflection cards. The cards were thought to become operative as a complementary tool to existing decision aids and information brochures that

tend to be fairly functional and contain a lot of rigorously collated information. The aim for the cards was to facilitate 'translating' such medical information meaningfully to one's personal life as well as possibly stimulating an interest in reading those brochures in the first place.¹⁰⁴ In this way, it was also hoped that the cards would support the role of doctors in assuming more of a facilitating role as part of the shared decision-making process by contextualising medical information and how this might be relevant in specific aspects of a screening participant's life. Furthermore, it was hoped that the cards would generate a useful ground to evoke and discuss preferences and, thus, factor them into the screening decision-making process accordingly.¹⁰⁵

Given these hoped-for uses, I chose cards as a lightweight approach in contrast to a multi-page brochure, as they can be used individually at home or when out of one's home. Furthermore, cards are already used for all sorts of purposes, such as games, learning and note-taking; hence the hope they would appear as a familiar object, unrelated to the medical context and something people would more easily engage with.

The question of how many cards should be developed rose early in the design and review process. It was specifically triggered as a consequence of phase one in the design and development process of what the cards should be and do. In this phase, I started out with focusing on developing scenarios which would allow people to get a sense of the various side-effects one may experience as part of the screening and treatment of prostate cancer. This seemed particularly important, given that these side-effects might have developed only as a consequence of having been overdiagnosed and overtreated (Welch and Black 2010), hence, theoretically at least, avoidable. Thus, I conceptualised the cards as 'task cards' describing an everyday social setting and scenario in which the card-reader should put himself in the coming days. The cards also ask the reader to respond to the scenario described within the actual situation.

¹⁰⁴ In my first meeting with Felix Huber, a GP and author of a carefully written information brochure on prostate cancer screening, he indicated that men typically read the information brochures only in the waiting room or briefly during the consultation, but less likely afterwards.

¹⁰⁵ Assessment criteria will be discussed in the following chapter.

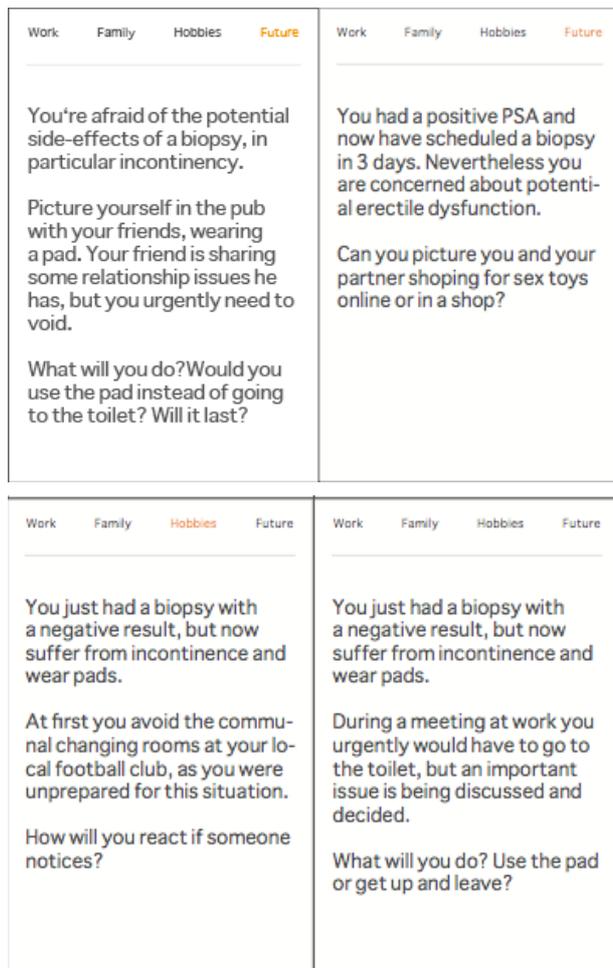


Figure 6.15: Prospective reflection cards - first prototype

I developed and reviewed an initial draft card with a number of people among family and friends, including one person who had been doing PSA tests for a number of years already. Furthermore, I also discussed the cards with supervisors as well as Eva Ebnöther, a professional medical writer who wrote the prostate cancer information brochure for the Swiss Cancer League. From a methodological perspective, and in the spirit of RTD, such early discussions were aimed at inflecting my designed intentions with and through the interpretation of the material by potentially relevant ‘users.’ This allowed me to become sensitive to the interpretative issues of using cards for the purposes I envisioned.

From the review and discussions, a few aspects surfaced as very problematic in this first draft. Firstly, both the design of the cards and the description of the scenarios seemed rather generic, and I had doubts about their capacity to evoke the desired reflections and responses. Secondly, and more fundamentally, the scenarios focused almost exclusively on handling social situations with side-effects. From the epidemiological data, there is also obviously the likelihood that no such implications surface. These situations

should equally be anticipated and become part of the card set. More generally, this triggered the question of what the balance of scenarios can and should be.

Lastly, and to my surprise, from my first discussion with urologist Cédric Poyet I learned that impotency and incontinence side-effects are experienced as a result of actual prostate cancer treatment and not of having a biopsy. This medical aspect was clearly wrongly understood and, thus, misrepresented in the cards. After learning this, I reviewed the NHS decision aid again and was deeply concerned about their very ambiguous description on this matter, which had misled me. In summary, for the next phases of development I decided that more focus should be placed on a narrative style of scenario, on fear and uncertainty of getting cancer instead of the side-effects of preventive treatment (as suggested by urologist Cédric Poyet) as well as deliberating and defining the underlying logic for the distribution of the social and medical scenarios across the card set.

Phase 2 - Size of card set and scenario distribution

I started this phase by considering on what grounds and logic the screening scenarios covered in the card set would be distributed. By this I refer to the practical question of how to decide how many cards there should be that help in reflection on a scenario where someone conducting cancer screening experiences side-effects, as opposed to the number of cards with scenarios where such negative experiences are absent. Producing a set of cards that only cover one type of scenario, such as when people screening for cancer experience side-effects, clearly deviates from the statistical reality of cancer screening (not everyone experiences side-effects), and thus may be said to be politically “biased”. Such a bias seemed unwarranted and undesirable as well as being potentially reminiscent of “fear-mongering”. Unbiasing the card set thus required clarifying and developing an underlying set of assumptions based on which the number of cards would be meaningfully defined and their content structured accordingly. So as to define this underlying set of assumptions, I mapped the prostate cancer screening process from taking a PSA test through to actual prostate cancer diagnosis and treatment. I augmented the map with the relevant statistical data that is also used, for example, in the NHS decision aid for a PSA test. This map shows the statistical distribution of potential scenarios that might affect a person during the cancer screening process.

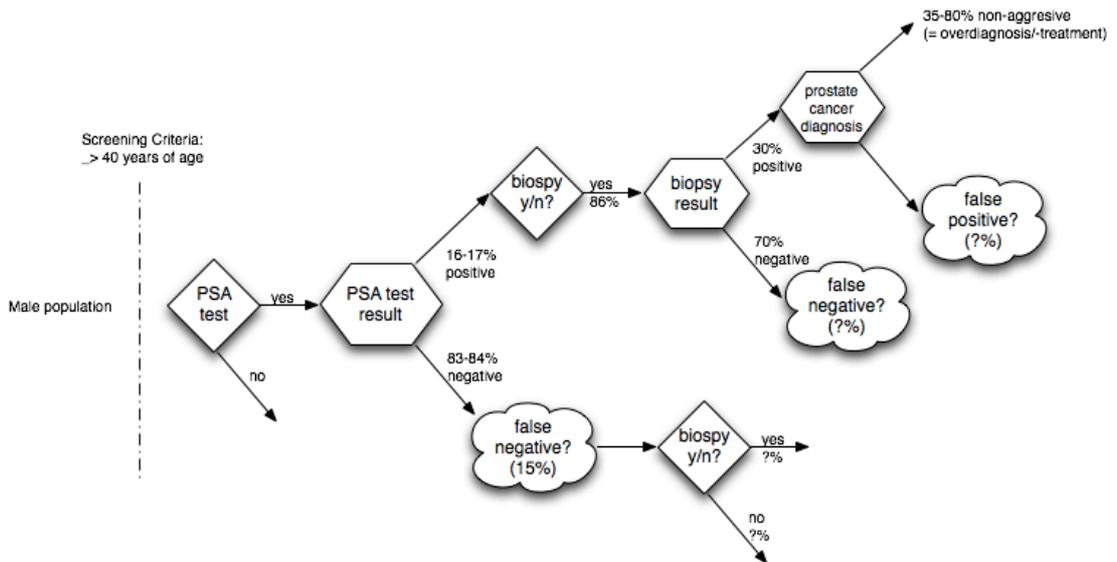


Figure 6.16: Prostate cancer screening process and informational gaps

This allowed me to do two things: firstly, project the total number of cards it would take should I decide that the whole screening process would be covered. Secondly, and more importantly, this map allows to reflect on the distribution and relative share of cards for specific aspects of the screening process, and what their content would likely need to be. As it turned out, settling these questions was fundamental to making the card set both intelligible and effective. This is best explained with the following example.

My initial assumption was to include the whole prostate cancer screening process. In order to define the total number of cards needed, one needs to start at the end of the screening process by determining the number of cards to be included for actual prostate cancer diagnosis. If we say there are 10 cards based on which the outcome of a biopsy would allow to stipulate a diagnosis, the distribution according to the diagram above would be as follows: 7 negative biopsy results (i.e. false-positive PSA tests) and 3 positive biopsy results, of which 1-2 would be non-aggressive prostate cancer largely depending upon age of the screening participant. Thus, regressing from these 10 cards that represent the 86% deciding to have a biopsy translates into 11.6 cards for that 16/17% having a raised PSA level, which incorporates those 10 cards. The 11.6 cards which represent the 16/17% with a raised PSA level, in turn, are a fraction of all people taking a PSA test, which would amount to 70.3 cards.

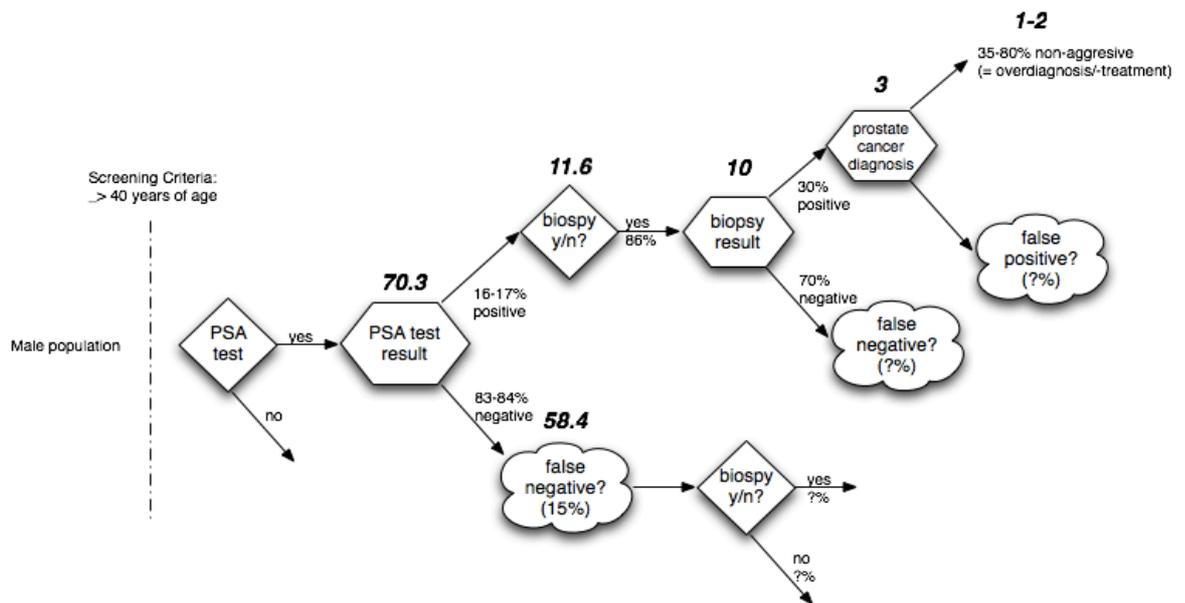


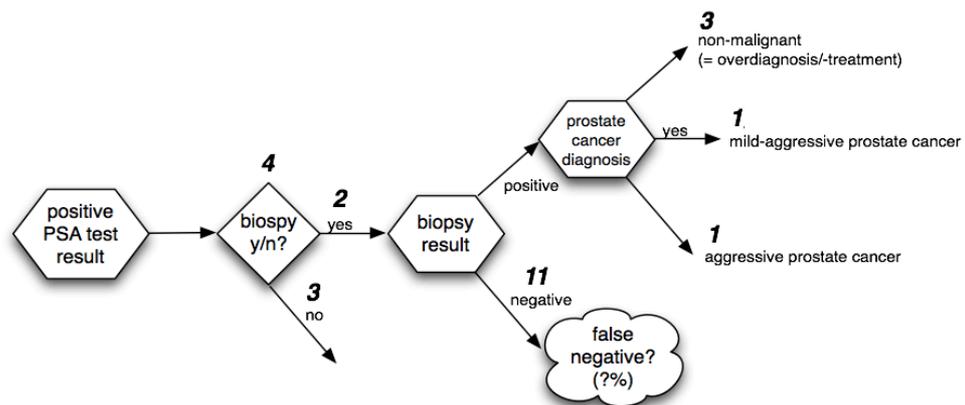
Figure 6.17: Projection of reflection cards for decision-making process

Needless to say, 70 cards seemed a lot both in terms of people using them but also in creating a meaningful set of them. In order to assess this, I deliberated what kind of scenarios and content would need to be created for the cards so that they would be meaningful for the phase of the corresponding screening process. With reference to the diagram above, this exercise included thinking about the content of 58 (70-12) cards for people with a negative PSA test result. The content for such a scenario would have needed to include the potential worry of having a false-negative result (15% of these results, source: NHS decision aid). Unfortunately, it is not known how many people worry about false-negative test results. Even if this figure were as high as 50%, this would have left approximately 30 cards stating that one has had a negative PSA test result without much further narrative potential. This seemed a comparatively high number in relation to the total set, thus risking rendering the whole set rather dull.

As a thinking exercise, I explored the idea of focusing the card set on the assumption that one has had a positive PSA test result or, in terms of prospective reflection, one imagined what one would do in the case one had one. The reason for this was also that, from the statistical information available, it clearly seemed that at least a great many people are seriously worried about a positive PSA test result. Among the population with a positive PSA test result, 86% undertake a biopsy despite the fact that 70-80% of these results are false-positive ones. A similarly high figure seems unlikely for people undertaking a biopsy after having a negative PSA test result, even though this figure is

unfortunately unknown. In doing this thought exercise, the 16/17% with a positive PSA test result transpired as the starting point for the cascade that would encompass all subsequent screening process phases and cards. The distribution logic for the cards is visualised in the following diagram. Adding up the number of cards for the various scenarios amounts to a total of 25 cards.

Distribution of cards



Distribution of cards

25 cards in Total

- 4 undecided about biopsy
- 3 decided not to have biopsy
- 2 decided to have biopsy, date fixed
- 11 with a negative biopsy result
- 3 with a positive biopsy result, non-malignant prostate cancer
- 1 with a positive biopsy result, mild-aggressive prostate cancer
- 1 with a positive biopsy result, aggressive prostate cancer

Reference statistical distribution:
BMJ Group decision aid

Figure 6.18: Distribution of cards under assumption of positive PSA test result

As a starting point for the card set, venturing off from a positive PSA test result seemed to make a lot of sense. The next question to be answered, however, was how far the cards should cover the screening process. More specifically, should they cover the phase of cancer diagnosis as well as cancer treatment, including the corresponding side-effects following cancer treatment? I decided not to include the actual cancer treatment phase, as this would have nothing to do with uncertainty as part of the anticipatory screening process itself. The latter ends upon cancer diagnosis and, thus, should theoretically clarify any uncertainty issues. I write 'theoretically' as the issues of cancer overdiagnosis and overtreatment obviously do leave significant degrees of uncertainty unresolved (Moynihan et al. 2012). But the character of uncertainty in this area has a different quality compared to the PSA test

uncertainty, and, thus, requires a different engagement with it.¹⁰⁶ This would have meant entering a field of medical expertise which I did not have and could not easily acquire nor credibly or ethically feel equipped to engage with. Most likely, I expected that engaging with these advanced issues of uncertainty would also potentially cross a boundary with the medical professionals that would not be favourable to my project, at the very least not in the initial phase of collaboration.

A further consideration was that the scenarios in the cards should reflect people in different stages of the screening process. This led me to the decision to have 4 cards for scenarios such as ‘people undecided about biopsy’. Typically, when looking at decision aids, such aspects of being in the process of deciding are not reflected. Instead, what is highlighted is the final statistical distribution of how screening eventually plays out across a population. With this, the focus for the cards in terms of which aspects of the screening process to consider was finally set.

Phase 3 - narrative style and focusing on uncertainty

Having solved the question concerning the distribution logic of the scenarios in the card set, I next focused on narrative style and putting more emphasis on the uncertainty of developing cancer instead of the side-effects from treatment of cancer.¹⁰⁷ Following the envisioned use and contexts, I also explored various form factors and paper qualities. Especially length of text used, size of font but also having enough space for taking notes would largely determine the size of the cards.

Furthermore, I aimed for a minimalistic design in order to make the cards visually light as well as inviting to look at. To achieve lightness, I got rid of the header separation line from the initial card design both on the front and back. I changed the font to what seemed a less formal looking one, the colour of the paper from white to ecru and a thicker paper with more texture. On the back, I exchanged the header ‘notes’ with various prompts to take notes that were related to the scenario on the front, such as “At first, I thought...”

¹⁰⁶ This is because uncertainty with overdiagnosis relates to the difficult question of prognosis of the pathological development of an identified cancerous lesion and what an adequate course of action may be.

¹⁰⁷ As suggested in the first meeting with urologist Cédric Poyet (see above).

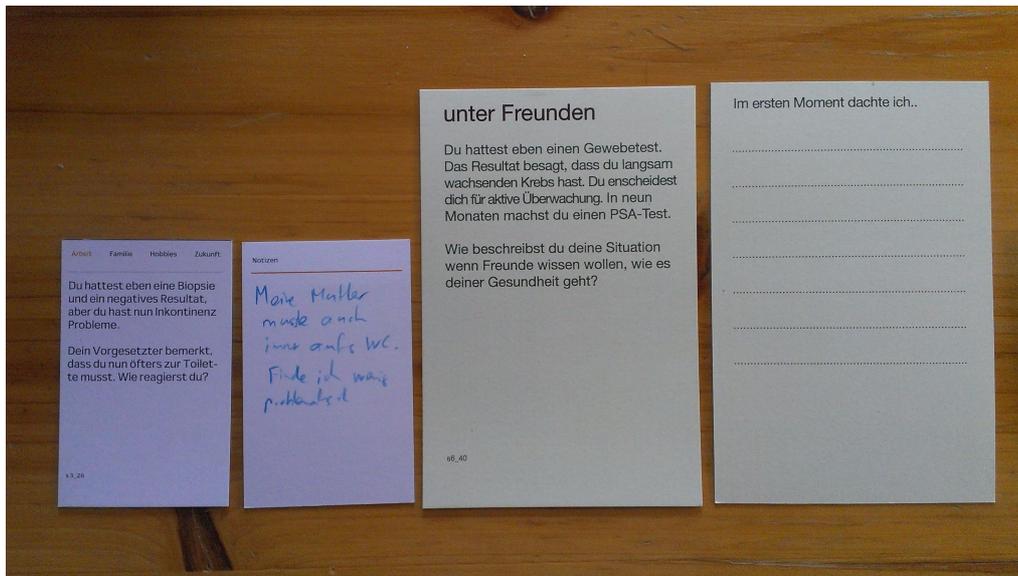


Figure 6.19: Early card prototype and updated one



Figure 6.20: Different sizes of card prototypes

(Translation: Love – The screening test result is slightly above the recommended threshold. Your GP recommends you to conduct a biopsy, just to be sure. How will you respond?)

The final structure of the cards was made up of a header section, the actual scenario, questions for reflection and, on the reverse, a prompt to take notes semantically related to the scenario described on the front. As a move towards a more narrative style, I discarded the generic category headers ‘work, family, hobbies, future’ I had in the first draft and started to be more context-specific. I tried to use the header as a way to describe the contextual setting of the scenario which follows. Examples used were ‘among friends’ or ‘you and your GP talk about next steps.’ This already seemed to relax the cards in their appearance and address. In trying to develop more narrative

scenarios, I aimed at shifting the focus on various imagined social settings in which a person engaged in screening could find himself. Particularly, I tried to generate scenarios where the question of uncertainty would be addressed socially, see for example below.

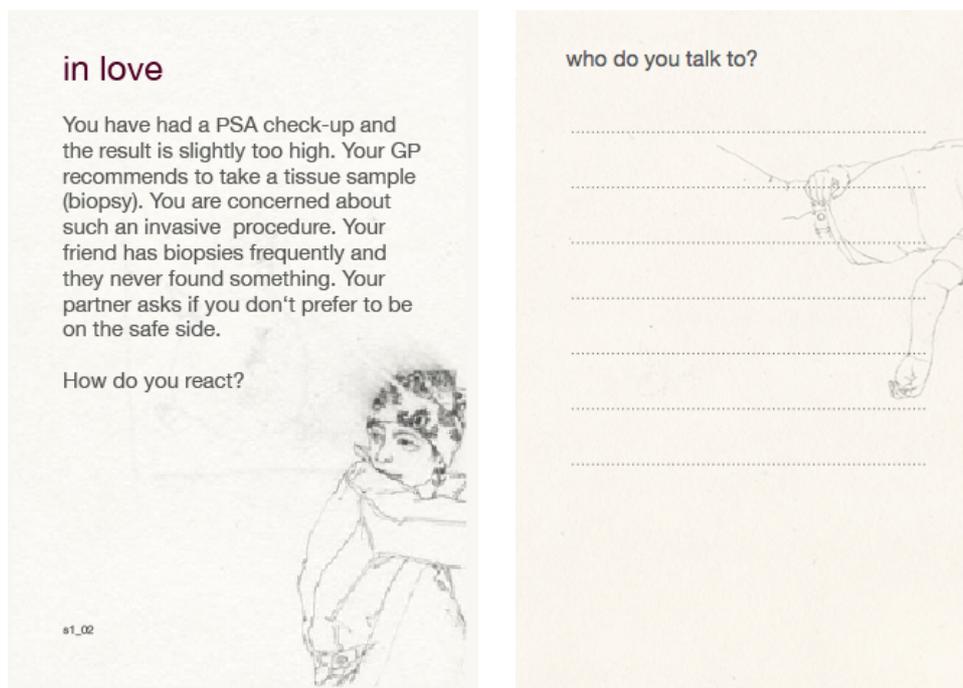


Figure 6.21: Reflection card final prototype, back and front

Review and discussion of the card prototyping process

Developing 25 cards with what felt could be relevant and realistic scenarios was challenging. As part of the scenario development process, I considered to open up this part of the process for future versions by engaging people who conduct screening and live through these issues of uncertainty. The way this could be realised was by running workshops or even, at a later stage, setting up an online version of the cards which would allow people to generate their own cards and share them with others. As a starting ground, however, I drew indirect inspiration from reading through relevant online fora entries of people engaged in prostate cancer screening.

Concerning the language used when developing the cards, I tried again to follow a degree of underspecification. This seemed difficult to achieve and I felt it ran the risk of the cards seeming technically incorrect or unprofessional by being too vague. Nevertheless, I tried achieving a subtle balance, for example by not using any quantitative probabilistic information so often found in medical decision aids and information brochures. Instead, I resorted to using subjunctive forms or at best descriptive words such as 'often,' or 'in

some cases.’ I followed this approach when describing the results of a PSA test. To give an example, a card would start with ‘You had a PSA test and the result was slightly below/above the recommended threshold.’

The final aspect I changed in the cards’ appearance, apart from size, paper texture and font, was that I decided to include illustrations. The reason for doing so was, again, to make the cards appear less sterile, which seemed important given their content and the hoped-for engagement with them. This was an indication I noticed during my various discussions when reviewing the prototypes. However, and importantly, what the illustrations should not do was to somehow represent anything specific that was described in the scenario on the cards. Rather, the hope was to further support evoking a reaction to or reflection on the card’s content. To that extent, the illustrations were selected based on their capacity for associative potential and, at best, to create a tension between the text and the illustration itself. With the integration of illustrations, the card set seemed finalised as a robust working prototype which should allow me to find out whether they are seen as valuable and productive in the eyes of relevant medical professionals. The full set of cards can be found in Appendix 3.

Conclusion

In summary, using research through design as a method for the design and development process was very productive in a variety of ways. First and foremost, it helped elicit a sense and understanding for the difficulty of designing for epistemic action. This challenge is made up of at least three aspects:

Firstly, the process of clarifying which particular informational aspect I should specifically focus on when designing for epistemic action and, thus, be able to support an understanding of. As I have highlighted, this requires a precise understanding of the context on the part of the designer, separating out aspects thought to be of lesser importance and formulating a working hypothesis as to what aspect is believed to be the key epistemic goal. As the design process described above has revealed, RTD allows for such a working hypothesis to be continuously refined. It does so by aiming to explore understandings in a particular situation, and, in that regard, has helped focusing the iterative design process described above towards supporting such a situational understanding. This, in turn, provided the lens through which I was able to prospectively reflect upon the relevancy of the potential interactions with the developed prototypes and refine them successively.

Secondly, the actual design of epistemic action requires identifying possible and relevant forms of interaction which may productively support reflection-in-action. Here, RTD was especially helpful in working through and reflecting-in-action various prototypes, in turn allowing to test and clarify understanding as well as evaluating the epistemic capacity of the designed interactions. Specifically, RTD helped clarify which role the users of my prototypes would likely take on when using them and the implications this might have for the desired epistemic action.

Thirdly, identifying the multiple possible and valid interpretations of any such epistemic action that seem to remain no matter how specific one has thought to have designed it. Importantly, what this does is supporting a (late) realisation of the potentially many valid ways in which text as well as visualisations can be read, interpreted and understood (Sengers and Gaver 2007). Given this, what is of importance is approaching the design of such decision support tools with a knowingness which permits “a condition of recognising the processing of the situation to unfold in a way that allows the user of the designed tool to also enter in some kind of composition with the design process” (Fuller 2014).¹⁰⁸ Shared decision-making (SDM), in turn, may also need to be understood as the actual *making* of the necessary conditions for a decision to be *made*, rather than today’s design of SDM that seems more akin to shared decision *taking based on facts presented*. With this, I specifically refer to the complexity and richness of the design process of such conditions that materialise in a decision-support tool as developed and described in this chapter.

The design and development process also evoked thinking about relevant aspects and questions concerning the evaluation. The expected outcome from the evaluation will be discussed at the start of the next chapter and followed by the actual findings from the evaluation and a discussion thereof. The latter will also reflect on the issues, methodological limitations and lessons coming out of the design and evaluation process.

¹⁰⁸ Personal communication, 02. 06. 2014.

Chapter 7

Evaluation and research findings

Introduction

This chapter will discuss the research findings. I will do so firstly by reflecting on the expectations and the hoped-for outcomes of the evaluation phase that emerged out of theoretical arguments in chapter 1 to 4 as well as the design process itself. Secondly, I will discuss the evaluation process and outcomes. Based on this I will then discuss and methodologically interpret the research findings. Lastly, I will highlight and reflect upon both methodological challenges and insights during the design and evaluation process.

Before proceeding to the discussion of the evaluation sessions and interpretation of research findings, it is useful and necessary to situate the hopes and expectations from this phase of the research. Together with the methodological interpretative framework discussed later, this will allow me to clarify both the role and meaning of the prototypical tools for this research as it emerged through the design process and evaluation sessions.

To start with, it is important to characterise the context within which this research seeks to make a relevant contribution. This research context can be described as a technically specialised field. Consequently, high barriers could potentially exist in terms of being accepted, in particular as a non-medical professional. This is because I am designing material which is to be used in the process of patient consultations. Typically, such material seems to be developed by medical professionals with an interest in decision support.

Furthermore, it has to be noted that I have no prior documented practical engagement or expertise in this area, thus the research is almost purely informed by a critical reading of relevant literature as well as my design practice in other non-medical related areas. It goes without saying that the opportunity to observe processes in medical consultations requires access to these contexts, which is not easily facilitated and involves a variety of ethical and professional approvals. When introducing myself with a background in design and cultural studies to medical professionals I spoke to during this research, I was met with interest; nevertheless, some explanation was required as to how I arrived at this topic and why it makes sense to engage with it from my background.

As regards more formal aspects, this research project is completely self-initiated as well as self-funded. All these aspects can and have made it difficult to get access to and cooperation with relevant medical professionals.

The process of identifying and winning over relevant medical professionals to cooperate with was very lengthy and ran in parallel to the design process itself as will be explained below. The evaluation sessions with the medical professionals must, thus, be seen from multiple perspectives.

As I have come to realise throughout the design and evaluation process, this constitutes the first phase of a much longer research project, which from a methodological standpoint aims at entering a more intensive co-creation process with relevant medical professionals and patients as well as eventually evaluating the resulting designs with both of them. Needless to say, such a project would operate on a scale and timeline that would require more substantive resources and infrastructures, rendering it well beyond the capacity frame of this PhD research. This, however, will only become possible once I have established and documented that there is a need and perceived relevancy on the side of medical professionals, and that my knowledge and experience are accepted and seen as desirable and complimentary within this context.

Following these contextual considerations, the desired outcome of the evaluation phase of this research project concerns three broad aspects. Firstly, I was hoping to get a sense of the topical relevancy of what I do with my decision support tools and, as a consequence, commitment by the medical professionals to cooperate short-term with the prototype evaluation. This was thought important so as to understand whether what I perceived as the decisional problem from the literature review maps onto the actual clinical decisional problems as encountered by physicians and patients alike.

Furthermore, prior research suggests (Bates et al. 2003) that the acceptance of decision-support tools is critically dependent on medical professionals finding the tools relevant to their consultation practice. Based on this as well as the self-funded nature of this research, I decided to only work with medical professionals in the first phase. This led me to contact relevant medical professionals during the design process phase of the project in order to clarify medical questions and ensure clinical relevancy of my research and design focus. Additionally, the hope was that topical relevancy and acceptance, in turn, might open the path to undertake field research concerning physician-patient consultations in the future. As a result of the decision to work with medical professionals but not patients in this phase of the research, no ethics statement is necessary.

Secondly, I was hoping to generate interest in the particular design approach towards decision support I chose. Feedback on this aspect was deemed crucial so as to better understand and steer relevant and useful future design process and prototyping directions. This aspect was factored into the

design process by following a design approach which aimed at generating prototypes which are minimal enough to invite participation in refining them, while at the same time being finished enough so as to be perceived relevant and productive for potential patient evaluation.

This latter aspect relates to the third and last hope of the evaluation phase. This is the perceived relevancy of the prototypes so as to generate interest and commitment for a longer-term collaboration with the goal to evaluate the prototypes for their relevancy, capacity to support the specific decisional aspects and potential in an evaluation with actual patients. With this I refer particularly to the potential gap between representations of the decisional problems in the literature, predominantly written by medical professionals, and accounts from a patient's perspective, such as those provided more recently by Kukla (2007, 2005) and Carel (2008).

In order to answer these questions, I organised evaluation sessions with three medical professionals. Professor Jacques Cornuz is a specialist in social and preventive medicine, highly engaged in the practice of shared decision making and director of the university hospital in Lausanne, Switzerland. I had identified him through a paper he co-published on the state of SDM in Switzerland (Cornuz et al. 2011). More specifically, I was interested in his appraisal of my tools and, in particular, their underlying design approach which is vastly different from existing decision aids. Given his broad experience and practical engagement in the field of SDM, his specific expertise was deemed highly relevant for my research and its application within the Swiss medical context.

Felix Huber is a GP with a long-term critical engagement with the various issues of prostate cancer screening described earlier. He is also the founder and president of a network of group practices of GP's in Switzerland. I had become aware of him through what I deemed to be a very balanced medical information brochure concerning prostate cancer that he had co-written (Huber et al. 2012). I vaguely expected that, in his professional role as a GP, he would be slightly differently involved in the role of advising people on whether or not to screen for prostate cancer compared to a urologist. The latter, I understood, would be someone to be involved once there was actual concern of potential prostate cancer based on a positive PSA test result with a GP patient. This question of the role of different professionals in the screening consultation process was something to be further clarified as part of the sessions, particularly since the literature was not informative on this aspect.

Lastly, Cédric Poyet is a urological medical specialist and a Deputy Head of the Prostate Carcinoma Center at the University Hospital in Zurich, Switzerland. In his role, he could be using my tool as well as engaging with

potential patient users on a daily basis. Thus, he could be said to have extensive relevant experience to comment on the challenges with prostate cancer screening in everyday medical consultation practice. This, I expected, would allow him to comment on and assess the relevancy and capacity of my prototypes from a variety of important perspectives: firstly, evaluating my tools from a medical technical perspective, which I thought was important so as to ensure they were technically correct and fit for purpose; secondly, his central position in matters of prostate cancer screening should enable him to give feedback on whether the tools focus on aspects relevant for better supporting the consultation process; lastly, his extensive experience would allow him to estimate the tool's capacity for epistemic action for a variety of different types of persons, be they young or old and with different visual and linguistic literacies.

Evaluating with medical professionals of different kinds was intentional. My aim was to generate multiple perspectives on the tools and the role they might play in different settings and for different medical professionals. This aspect seemed particularly important, given the absence of practical insights into the consultation practice in these settings. In organising the evaluation sessions, I was able to meet Cédric Poyet and Felix Huber once before the actual sessions to personally introduce myself, this research and my interest to win them as participants for the evaluation sessions. With Jacques Cornuz I was only able to have e-mail exchanges in which I introduced the research and he declared interest in participating in such an evaluation session.

There were also much more immediate hopes and expectations concerning feedback on the prototypes and their potential role as part of medical consultations. One role the prototypes were hoped to play was that of a useful inflection point for discussing the problem space with medical professionals. I was particularly interested in learning how they assess the project based on their everyday medical practice and the kinds of challenges they meet in explaining prostate cancer screening and its potential implications. Thus, inherent was the question of which aspects were to be supported in the medical consultation process and should, thus, be focused on in the prototype design process.

Given the difficulty of accessing relevant medical professionals early in the design process, a problem that can be easily imagined to surface in a variety of contexts and, more particularly, in the context of entering a new problem space and field, the design focus decisions I took were exclusively based on my review of the literature. To some degree, thus, I had to speculate on whether such literature represents well the kinds of challenges medical professionals would deem (most) relevant.

Directly following questions of what to focus on were also considerations of how the tools would specifically be put to use as part of the consultation process itself. This was an area which seemed most difficult to prospectively assess, because of the previously mentioned absence of prior experience with the everyday medical consultation process. As discussed in the previous chapter, the tools were designed so as to be used by actual decision-taking persons instead of something to be operated by the medical professional. The question, thus, was whether the medical professionals would see this as both desirable and well supported with the tools I created.

Inherent in how the tools would become operative in the consultation process was the research question of whether the medical professionals would see the potential and capacity for epistemic action in the design of the tools. Having medical professionals recognise a capacity in the tools for clarifying false-positive and false-negative results as well as the preference-construction and diagnosis process would indicate that the tools a) facilitate translating medical information from DA's to one's personal life, and b) allow doctors to take on more of a facilitating role. In turn, this would imply that screening decisions would be less often deferred from patient to doctors and more often taken in a shared decision-making manner.

Evaluation sessions

The sessions were planned as semi-structured interviews along the following lines: First, I gave an overview of the evaluation session and what would happen within the planned total time of about 45 minutes. This is the amount of time I deemed appropriate when asking medical professionals to participate, and long enough to achieve the engaged discussion I was aiming for. Then, I gave a brief introduction of the screening decision problems, my analysis of the shortcomings of existing approaches and an overview of the tools I designed. Subsequently, I introduced the Venn diagram interaction tool, explained how it works and highlighted the specific decision problem it seeks to address. I explained the tool by way of a scenario when a patient has undertaken a PSA test and received a result of 3.7 or 4.1 score. Additionally, I explained the particular inspiration for the design stemming from an EMT perspective as discussed previously.

Following this introduction I started the conversation by asking the following questions:

- Is the tool understandable for the medical professional and patient?

- What is your assessment of the capacity of the tool to support clarification of the decisional problem relating to understanding false-positive and false-negative test results and whether or not to have a biopsy?
- Will the tool be deemed relevant and used by decision-takers?

After having explored the specific questions concerning the Venn diagram interaction tool, I introduced the reflection cards, how they work and the specific decision problem they aim at addressing. This was followed by highlighting the particular sources of design inspiration I gained from EMT and by starting to ask the following questions:

- Are the cards understandable for the medical professional and patient?
- What is your assessment of the capacity of the cards to support clarification of preference diagnosis and, thus, whether or not to engage in prostate cancer screening?
- Will the cards be deemed relevant and used by decision-takers?

After having explored the tool #2 specific questions, I directed the discussion towards more fundamental questions concerning both tools and their overall role in such SDM processes. The questions I had were the following:

- Was the selection of thematic focus relevant for supporting SDM processes?
- What will happen if tools were used by patients?
- Would you use the tools yourself? What would you want to change/improve?
- Do they make good use of doctor time?
- Do you have an interest in further collaboration for tool refinement and evaluation?

Session with Jacques Cornuz

The first session was with Jacques Cornuz and took place at the University Hospital in Lausanne on the 21st of February 2014. I would have preferred to have this session last, since this was the first time we met, but the availability of doctors was beyond my control. Thus, I anticipated needing a little more time to introduce myself as well as the research project. However, given his expertise and engagement with SDM, I thought there was not too much of a risk losing much session time on this aspect. Hence, I approached this first session optimistically and with great curiosity.

When I met Jacques Cornuz in his office, he quickly asked his colleague

Reto Auer, a medical doctor with a background in epidemiology and biostatistics and Head of the clinic at the university hospital, to join as he would have to leave after 45 minutes to attend a funeral. This seemed to me a 'good sign,' indicating that my research was of interest to them and they were anticipating a longer conversation. I started the session by introducing myself in terms of professional and academic background.¹⁰⁹ I then introduced my academic research project, which - in a nutshell - ventures from the observation of an online health search phenomenon and then builds on a critical review of evidence-based medicine and existing decision-support approaches in SDM with the goal to re-design decision-support tools inspired by extended and distributed cognitive theories.

Interestingly, both of them pointed out quickly that SDM does not really work well today for a variety of reasons, key among them being the lack of interest on the side of patients/participants. This candid acknowledgement was partly surprising to me, as the literature pointed more towards a problem of adoption by medical professionals. However, my very own appraisal of existing decision-support approaches (see chapter three), hence the motivation for this research, was that there was a 'design problem,' even though I was not aware of the magnitude of the lack of interest by participants to this point.

After a brief discussion on the decision problems involved in prostate cancer screening, I introduced the reflection cards by explaining how they should be used and could be integrated in a consultative process. I also explained the thinking behind their design and the hope they could support extended and distributed cognitive processes. As I expected this aspect to be of much less interest to the medical professionals, I kept this part rather short when preparing for the evaluation sessions. Much to my surprise, this turned out to be of substantial interest to both Jacques Cornuz and Reto Auer. They explained that, as part of their engagement in the development of shared decision-making from a perspective of preventive medicine, they have witnessed a number of other tested and failed approaches to facilitate behaviour change. Among them was for example behavioural economics which, for them, was incapable of tackling their professional challenges, such as getting pregnant women to stop smoking.

The cards were first met with what seemed reserved curiosity by the medical professionals. One of the first questions from Jacques Cornuz was whether the cards operated with precise quantitative risk information as part of the scenarios. I explained that they did not and that this was one of the

¹⁰⁹ Professionally I explained that I am a lecturer and deputy Head of Interaction Design at Zurich University of the Arts, as well as having 20 years of professional experience in various roles as a designer, consultant, researcher and analyst.

intentions in designing cards. Furthermore, I explicated that the purpose of this linguistic underspecification was, on the one hand, to stress the complementary function of the cards to the more extensively informative brochures and, more particularly, the consultation with the medical professional; on the other hand, to avoid the impression of an exact figure which serves as a reliable ground for planning and individual decision-making in this specific context. In relation to the latter, the absence of such quantitative risk information should, in the best of cases, direct attention to the prospective reflection facilitated by the questions on the cards.

The two medical professionals agreed that this was an interesting take on designing decision support, in particular also with regard to respecting the literacy levels of some of their patients, which can be problematic with extensive information brochures that include such quantitative risk information. Furthermore, they noted one aspect of the reflection cards which was structurally very different from existing decision-support approaches. For them, what the cards did was introducing and, in this way, reframing the time dimension of decision taking. Existing tools seemed designed for something akin to instant decision taking or, as they put it, “sitting down, decide and then execute decision” based on the current life of a decision taker (Kukla 2005, 35).

The reflection cards, in contrast, introduced a much longer time dimension both for the actual decision-taking process and the time-frame for which the decision should be accounted for. In other words, the decisional goal supported by existing tools is to make the best decision for the current situation of the decision taker, and the information provided seems to be selected and designed particularly to support this motivation. My approach, in contrast, aimed at both factoring in a longer time-frame for which the decision taken should be valid and, as a consequence, explicitly anticipating shifting values and criteria for assessing and accounting for the decision.

Given the long time-frame a prostate screening process occupies (up to 25 years), such an approach seemed very interesting and valuable to them. This perspective was inherent in my design approach, especially as a consequence of the inspiration found in Esposito's future present (2007). However, it was less apparent to me that the orientation of current decision aids seemed designed so clearly from and for a perspective of current lives only. In retrospect, this is much more obvious now. Since decision aids are not yet inclined to also support the much longer-term and potentially time-consuming preference-construction process, their focus is on the short-term preventive action avoiding the long-term significant risk of cancer.

Building on this observation, we also discussed the potential of a digital

version of the cards. My idea for an online version was that it should make accessible and visible the notes and decisions of other prostate cancer screening decision takers. This, in turn, would allow other people (screening decision takers as well as related ones) to understand and potentially be inspired by the ways in which such circumstances and decisions are being taken and accounted for. The potential benefits highlighted by Cornuz and Auer would also be to make the decision process a less isolated one as well as making visible that it is less linear than one might think.

This latter aspect is concerned with the experience of learning what it means to be doing screening. As it transpired from my literature review and was confirmed in my session with Cédric Poyet later, it is quite common for a person engaged in prostate cancer screening to have multiple biopsies based on false-positive PSA tests. Thus, it is not unlikely that such an experience might change people's minds as to the value of PSA testing and undertaking biopsies. To this end, both Cornuz and Auer saw the cards as an interesting and lightweight approach that might be productive for a variety of people and contexts.

However, they also pointed out two likely limitations for the productivity of the cards. The first is that, based on their experience, men are much less likely to actively engage in a shared medical decision-making process. As they explained, typically men will ask the doctor for a recommendation of what to do and generally follow this suggested course of action. According to Cornuz and Auer, men within European western culture tend to be less inclined to engage with their health and body.

Secondly, Auer remarked that, while he sees potential in the cards, they will not be able to 'unbias' the medical professionals themselves. With that he meant that, often, medical professionals have a relatively set opinion as to the best path of action and stick to that. Thus, as discussed earlier, the source of the problem for acknowledging preference-sensitive care and preference diagnosis would, according to Auer, remain unchallenged and unchanged. This was an interesting comment which I did not anticipate, as my working assumption was that doctors may have an interest in SDM so as to redistribute the responsibility for implications of the actual course of action followed. In retrospect, it may be said to be only partially surprising to hear such a comment from a medical professional engaged in the area of SDM: unbiasing medical professionals could be seen as one of the core problems to address since the identification of unwarranted small-area variation by Wennberg (1973).

In the second part of the session, I introduced the PSA Venn interaction tool along with the decisional problem it seeks to support. Interestingly, while

Cornuz immediately understood the tool and found it useful, Auer had problems in making sense of it. He was particularly irritated by the visual form chosen, given that in his medical training he was not used to a Venn diagram but to a line-graph for such matters. While the visual form did not work for both, the tool was of interest to them as something to be handed over and handled by the patient after a brief introduction. Also, they found the topical focus of the tool - enabling the shift from an epidemiological perspective to an individual one - very relevant for SDM. In the case of Auer, who first questioned the need for patients to understand this aspect, his opinion shifted after having him imagine being a patient having had a PSA test and then using the tool to make sense of it.

Another question the two doctors had was whether and how I saw my tools to complement existing decision-supporting material, such as brochures and videos. The question probably stems from the fact that some of this material is already in circulation and, thus, any new material should add to rather than duplicate that. My response to their question was that a key element that informed the design of my tools was the idea and hope that particularly the reflection cards would precisely facilitate the process of appropriating technical information (as provided by brochures and videos) by transferring it to a person's personal and social life and, thus, making it socially meaningful.

They both agreed that the tools were interesting for their purposes and agreed to proceed with the project to a patient evaluation phase at their hospital. Interestingly, with regard to my question of what they thought the reactions and outcomes of actual patients using the tools would be, they responded that they simply did not know, as my approach was structurally so different from previous ones. Lastly, they saw a great need to include non-medical professionals with backgrounds such as myself into the process of further developing SDM-tools, and invited me to an SDM meeting in July which would bring together French and Canadian medical professionals engaged in the SDM field. The meeting would address SDM in a number of other cancer-related fields, such as colon cancer, which, according to Auer, is a somehow opposite problem when compared to prostate cancer. With colon cancer, treatment is very effective, but very few people engage in screening and can, thus, be diagnosed and treated.

Session with Felix Huber

The session with Felix Huber was held in his office in his group practice in

Zurich. Because we met once before, the session needed much less introduction. I briefly gave an overview of how I thought we could run the session, asked for permission to audio-record it and started introducing the reflection cards. In the first meeting with Huber, I had already shown him an earlier prototype of the cards, so he was vaguely familiar with them. Thus, in the evaluation session, I explained the phase of the screening process they are focused on and the distribution of the scenarios according to the epidemiological statistics so as to be unbiased.

His first impression of the cards was that they presuppose a relatively high level of understanding on the matter of prostate screening on behalf of the reader. For him, a patient would first have had to read an information brochure on prostate cancer, such as the one he wrote, that introduces the topic. Hence, he was sceptical about the cards suitability for daily consultative purposes. When asked where he saw the potential difficulties arising, he referred to his observation that few men read their information brochure about prostate cancer screening.

Furthermore, about two-thirds of the men that see him in this regard simply ask him to tell them what to do, without any desire to engage in the decision-making process. For him this is not a satisfying experience, as he would rather see the patient taking this decision himself. I explained that my hope was that the cards might generate a deeper interest in the matter by a) situating the potential course and implications of the screening process in actual life circumstances, and b) framing the decision process over the course of a week, which - in turn - might lead to more engaged conversations with a GP. Huber thought that this might indeed make patients more interested in the matter. Nevertheless he admitted that generally men as a target user group were very difficult to work with as they often seem to lack an interest in engaging with their health. He also thought about whether to make the language used in the cards even simpler so as to address more potential users, but simultaneously worried whether such a strategy would allow to still produce meaningful cards supporting the preference-construction process. He suspected that, whilst he found the tool very good, it might not be used very often for the reasons mentioned above.

At this point, the conversation took an unexpected turn. Huber suggested to focus on women and breast cancer, instead, and develop the very same reflection cards for this purpose. For him, the cards would be very likely to work well in this context, as the breast cancer topic is emotionally a lot more charged and, in contrast to prostate cancer, affects women early in their lives. He mentioned that he also wrote a corresponding information brochure for breast cancer and that, in the gynaecological practice of a relative, women

are requesting and using the brochures a lot. Adding to this was also the current news that the Swiss Medical Board, an institution with the power to recommend but not to oblige, had decided to no longer recommend the use of mammography as the potential benefits do not outweigh its disadvantages (2014). Huber also offered to put me in contact with said gynaecologist to work on the reflection cards for breast cancer. For him the issue that needed be understood by women is the fact that many are overdiagnosed and overtreated, which as a problem is similar to the prostate cancer screening issues discussed earlier.

In the second part of the session, I introduced the PSA Venn interaction tool. Huber immediately confirmed that the tool makes the dilemma of interpreting PSA test results very easily understandable visually through the overlapping of the health and pathology circles. For him, the ability of the tool to show that even a person with a PSA score of 8.0 can still be healthy while someone with a score of 2.0 may already be having cancer was very useful to support a consultation. Furthermore, the tool would make clear that only by conducting a biopsy could the shift towards an individual's perspective, or from overlapping to adjacent circles, be made. Also the possibility to add notes and use the tool over an extended period and possibly multiple PSA tests was desirable for Huber. While he was sceptical regarding the broad use of the reflection cards, he was nevertheless willing to use both tools with patients by setting up a corresponding patient evaluation project in the months ahead.

Session with Cédric Poyet

The final session with Cédric Poyet, the urologist, was held in my office at Zurich University of the Arts. I started the session by introducing the reflection cards, how they were developed and what part of the screening process they cover. Poyet read the scenarios of a few cards and then started to reflect on their potential from the perspective of his everyday consultation experience. The first thing he noted was that some of the scenarios might be more likely to occur and, thus, be relevant in a setting of a urological specialist rather than a GP. The card and scenario he referred to was where a person already had a second negative biopsy. From his perspective, such a scenario would be less likely to happen with a person being in consultation with a GP.

Furthermore, he reported that, from his experience, people in screening are less concerned with potential side-effects of cancer treatment should that ever be necessary, but much more with the uncertainty of whether or not they

have or will develop cancer. The uncertainty seems to even increase in the scenario just mentioned above where a person had undertaken multiple biopsies with negative results. In such situations, Poyet mentioned that people often find it difficult to accept that the biopsy was not able to find cancer despite the (false-) positive PSA test result and continue to think that the cancer must be there. It is almost as if there was some kind of anchoring or path-dependent thinking being triggered by those false-positive PSA test results.

I followed up with Poyet whether my imagined scenario that a person may have had multiple negative biopsies over the course of the screening is a common one. Poyet confirmed that it is particularly common at the Prostate Carcinoma Centre at the University Hospital, as very often they get referrals from external urologists asking them to undertake another biopsy so as to confirm their own biopsy results. I explained the reasons why I included such a scenario in the cards, which stem from my review of existing information brochures and decision aids. I noticed that none of the brochures and decision aids I had reviewed make this operative aspect of doing screening very clear by explaining that a person may find themselves with recurring positive PSA test results and, as a consequence, undertaking multiple biopsies which turn out to be negative.

Additionally, the reviewed information material does not highlight how this affects people's sense of uncertainty and the discomfort this generates. This clarification led to another very interesting question from Poyet, i.e. who I would hand these cards out to. At first I thought he meant this as a general question regarding who would distribute the cards, but he quickly clarified that his question was whether I intended to hand these cards out also to people that have already had a third negative biopsy, similarly to the scenario just discussed.

My original intention was that the cards should primarily be directed at people not having started with prostate cancer screening and trying to decide whether or not to do so, or people who just got their first PSA test result back and want to make sense of it. At the same time, my thinking was that I could obviously not prevent these cards from possibly finding their way to people already engaged in prostate cancer screening who experience uncertainty about its merits. Poyet was of the opinion that the cards would compel such people to grapple with this decision actively.

Responding to my question of whether people would actually use the cards and actively engage in the decision-making process, Poyet responded with a 'yes and no.' Like Huber, Cornuz and Auer, he thought that maybe a third of the people would use the reflection cards while many would also

simply delegate the decision-making to the doctor. Again, men turned out to be a particularly difficult group for such active involvement in health decisions and, again, women were suggested as more likely to be open to such a decision-support tool. For him, the best way to motivate people becoming involved in their health would be by introducing such tools as part of the consultation process, either by a doctor or a nurse.

When I asked Poyet whether he could see himself using the tools in his institution and consultation process, he explained that the problem might be that he does not see the people who are still at the stage of deciding whether or not to start prostate screening very much. These are more likely to be found in consultation with a GP practice. He usually sees people being referred by a GP with a heightened PSA based on which the GP requests a biopsy to be undertaken. In such a situation, he believes that the reflection cards could be counterproductive by reopening the question of whether or not to screen. This could potentially harm the credibility and trust in the GP on the part of the patient as well as possibly trust in medicine as such. In such a situation, Poyet explained, he would first do the medical examination as requested and only then possibly re-evaluate the person's medical situation and discuss it with him as a next step.

This was a rather unexpected moment in the evaluation session. The reflection cards seem to have uncovered an aspect of the social and professional understanding of and dynamic between various medical roles in which a person may find himself. To a degree, I was a bit shocked by the apparent and implicit shared understanding among medical professionals that seems to prevent a medical specialist from inquiring into a referred patient's full understanding of the potential implications of the screening process he is about to start. The reflection cards seemed to have uncovered this by way of situating the scenarios in a variety of different moments within the larger screening process, some of which - according to Poyet - belong more to the world of medical specialists like him, while others belong to the practice of GPs. Thus, by being very specific in terms of the scenarios, the cards had unintentionally operated as probes into the professional practices of medical professionals and the underlying understanding of their role within the larger screening process. Again, neither from my review of the literature nor that of decision aids and information brochures would a prostate cancer screening decision-taking person become possibly aware of such aspects. From my perspective, I would particularly expect the medical specialist to make me aware of the full benefits, implications and risks of the procedures he is likely to know much more about than a GP.

I asked Poyet whether he could imagine the following scenario: a

person with the age to potentially start with prostate cancer screening visits his GP; instead of doing a PSA test, the GP hands him the cards and refers him to a specialist, such as Poyet, for an initial consultation excluding any medical examination. Poyet confirmed that this would actually be his most desirable situation, as, based on his particular expertise and experience, he thought he is in a much stronger position to consult a person in this regard and answer their questions. For him, the worst scenario, and simultaneously probably the very reason why the PSA test has become so disreputable, is that the PSA test is just undertaken without prior reflection on the potential implications thereof by conducting a preference diagnosis with the person.

Given the strong interest Huber and Poyet share in having well-informed and actively involved people in the decision-making process, I suggested to set up a meeting between the two so that we could discuss a patient evaluation project of my tools which would also consider a change in the consultative process between the GP and the specialist as just laid out. Poyet was very interested and enthusiastic about it and even suggested that this could lead to setting up a regular prostate cancer consultation hour where not just people interested in the matter but also their GPs could participate. Poyet explained that they successfully run a similar model with oncologists when having to discuss whether or not certain patients should be put on chemotherapy. Because the patients already know Poyet well and do not want to consult a medical specialist totally unknown to them, Poyet asks the oncologists to join the consultation he has with the patient and so they discuss the questions the patient has in that environment.

Furthermore, for Poyet it would be important and interesting to see how GPs perceive the specialist's role and what they could contribute towards a better collaboration with the patient in mind. He feared that GPs may sometimes be hesitant to refer a patient to the specialist as GPs see them as too eager to conduct medical analysis, such as biopsies. Interestingly, some kind of self-reproducing feedback and feedforward loops appear to exist:

On the one hand, a specialist such as Poyet sees patients referred to them by GPs and external urologists; he believes he has to conduct the medical analysis expected of him as a specialist by the GPs and external urologists without having the freedom to independently consult those patients about the process. Doing so, he believes, would question the credibility of and trust in those external medical professionals on the part of the patient and, consequently, negatively affect the relationship between the medical professionals and the specialist.

On the other hand, Poyet fears that those external medical professionals see the specialists as too eager to conduct medical analysis only, and are

hesitant to refer patients because of that. It is almost as if there is a vacuum inhibiting an explicit clarification of what the expected and, more importantly, most suitable role is for a specialist, which seems to stem from an implicit understanding among those medical professionals.

I suggested to Poyet that the potential of having a regular prostate cancer consultation hour could also include the aspect that people could learn about the uncertainties involved in prostate cancer screening without first having to have a PSA test and, subsequently, finding themselves in a situation where they feel they have to continue with medical analysis in case the result is near or above the threshold. This is something he confirmed.

This borderline test result situation brought me to the second part of the evaluation session where I introduced the Venn interaction tool. Poyet's first question was who would be handling the tool and I explained that this would be the doctor so as to show to the patient how to use it. Poyet commented that this would be a useful visualisation to demonstrate the different risks, such as when someone never had a biopsy and, thus, the overlapping zone would be much larger. At that stage in the discussion he also showed me the Prostate Risk Calculator, an Apple iPhone application developed by the Prostate Cancer Research Foundation Rotterdam. What was particularly interesting is that the application makes a difference between detectable cancer risk and relevant cancer risk, which in the exemplary case was 10% for the former and 1% for the latter. The difference between the two lies at the heart of the problem of overdiagnosis and overtreatment (Esserman et al. 2013).

Poyet, like Huber, Cornuz and Auer, was very interested in evaluating the tools with actual patients. As discussed above, for him such an evaluation project would make most sense if it is set up in such a way so as to also consider and look into the referral processes between the various medical professionals typically involved in the screening process. I fully agreed that evaluating the tools would make most sense with such an extended understanding of the relevant environment and processes to be considered.

Summary of research findings, challenges and methodological interpretation

To summarise, the evaluation sessions were very productive and informative in a number of different ways. I will first summarise the key findings before proceeding to discuss their methodological interpretation as well as challenges I faced during the research. The next and final chapter will reflect on how to design and research with extended cognitive processes in mind by

way of research through design.

To start with, the three medical professionals who found the thematic focus of my tools relevant and were interested in participating in an evaluation session were all highly relevant and qualified to comment on my research from both a medico-technical and a clinical practice perspective. The fact that, at the end of the evaluation session, all three of them confirmed an interest to further collaborate and evaluate with patients is a strong general confirmation of both the topical relevancy and potency of the design approach chosen and developed in this research. Furthermore, and more specifically, the prototypes functioned well as an inflection point for discussing my understanding of the problem space which in turn materialised in the prototypes.

The Venn interaction diagram was immediately found particularly useful for clinical practice and thought to exhibit a high potential for clarifying with patients the issue of the probabilistic nature of PSA test results and the inherent potential of false-positive and false-negative test results. Very little introduction and explanation was necessary. Thus, with regards to the Venn interaction diagram, there was a good match in terms of the hypothesised thematic relevance and the actual clinical and decisional situation to be addressed and clarified with patients.

In terms of the relevancy of the reflection cards, I have come to a slightly different appraisal. The cards were generally found relevant, interesting, well-developed as a tool and, at least for some of the people in prostate cancer screening, suitable and acceptable. Their potential for epistemic action was much more difficult to assess by the medical professionals. Since my design approach is rather new within the domain of medical decision support, there simply seems to be no prior experience to which the reflection cards could be compared to and, thus, a potential patient response to be inferred from such experience by the medical professionals. Considering the mostly information retrieval-based design approach underlying contemporary decision-aids and lack of uptake in SDM, there seems generally very little ground to compare novel approaches building on more explorative involvement on the part of patients.

Simultaneously, and more fundamentally I suspect, at the heart of this inability to imagine potential responses of people to the reflection cards lies an absence of the practice of carrying out preference diagnoses, as has been argued by Mulley et al. recently (2012). My speculation was partly confirmed by Poyet, the urologist, who estimated that many a medical examination would no longer be necessary if a preference diagnosis and careful SDM were better supported before starting the prostate cancer screening process.¹¹⁰

¹¹⁰ Personal communication, 14. 07. 2014.

Nevertheless, and specifically from a SDM perspective as noted by Auer and Cornuz from the University Hospital Lausanne, of interest and value was in particular the ways in which the cards reconceptualise the time-frame *within* which the decision should be taken as well as *for* which it should be valid and accounted for. The difference noted was that the cards, in contrast to existing information brochures and decision aids, view and, in turn, support decision-making not as something to be made instantly as part of the consultation with a doctor (Kukla 2005) but more like a process over the course of days or weeks and possibly a second consultation. Such a notion of time lies of course at the very heart of the extended mind thesis and distributed cognition, which, as discussed earlier, build on the premise that in some situations cognition is supported by the social and artifactual environment. In turn, such exploratory processes of cognition are best allowed some time to unfold.

Interestingly, while all of the medical professionals had an opinion as to how many patients might use the tools, none of them could imagine what would happen if the tools were being used by actual patients as part of the decision-making process. All of them were very interested in and affirmative of the underlying cognitive understanding based on which the tools were designed. This, I speculate, may have indirectly led to the perceived potential in the tools, hence the interest of the medical professionals to evaluate them with actual patients. The question of how these tools would support such extended cognitive processes with patients will, thus, be evaluated in a future research project.

Unexpectedly and importantly, my prototype tools came to function as probes for diagnosing the underlying and practiced understanding of and dynamics between medical professionals and their roles. This is particularly significant because the acceptance and productivity of SDM-tools, such as the ones I have designed, may depend on a subtle understanding of such dynamics as my literature review has shown (Bates et al. 2003). Thus, research through design as an approach seems to have been productive in non-intrusively unearthing ways in which designerly interventions may *be introduced as prototypes but simultaneously come to act as probes or, as I suggest calling them, **probotypes**.*

A brief recapitulation will help analytically clarify and position probotypes in relation to prototypes and cultural probes. At the outset, this research was influenced by two constraints which were actively and reflectively considered throughout the design process. The first one was the fact that, by and large, my problem space assessment and thematic focus for the probotypes were based on a literature review rather than any clinical

practice appraisal. The second one was the hypothesis of this research, i.e. that the decisional challenges are best seen and addressed as exploratory types of interactions conceptualised as extended cognitive processes. The synthesis of these two constraints led me to develop what I call a *minimal reserved* approach towards the development of the prototypes. This minimal reserved design approach can be characterised as follows:

Firstly, while I had a firm sense for the thematic problems to be focused on, I had no way of knowing how these were actually dealt with in clinical practice. In turn, I kept the design firmly focused on supporting epistemic action, while keeping explanatory embedding of the design to a minimum. As can be seen, there are no instructions on how to use either of the tools or, in the case of the Venn interaction diagram, any explanatory notes on the tool itself.

Secondly, and with regards to EMT, by way of the design process I successively came to realise the many possible interpretations of the epistemic actions I designed, no matter how detailed and focused I believe them to have been designed. This, in turn, affected my view and expectation of the role of my prototypes in the evaluation phase. I would argue that prototypes are classically developed with a very clear use-case in mind, meaning that both the problem addressed and desirable outcome and the use-context are (thought to be) relatively clear. As a consequence, the evaluation of prototypes is usually aimed at confirming the effectiveness and efficiency with which the prototype achieves the desired goal.

Cultural probes may be conceptualised and employed as a way to explore a problem and design space by way of engaging participants with the probes. Gaver et al. (2004, 53) originally intended them to be “collections of evocative tasks meant to elicit inspirational responses from people - not comprehensive information from them.” In that sense, while they are directed towards exploring a thematic area and how this manifests in the daily lives of people, they are not functionally directed to address or solve a specific problem.¹¹¹ In turn, the evaluation of the material produced by cultural probes is focused on allowing designers to get a better sense for a thematic context from the perspective of affected and engaged people and, thus, empathically inspire design which “could enrich people’s lives in new and pleasurable ways” (ibid.).

Prototypes, as the name indicates, somewhat occupy a middle ground. They are semi-directed, by which I mean a) functionally directed to

¹¹¹ Cultural probes in the beginning were often somewhat misunderstood and overly rationalized, even to the extent of using them to do requirements analysis (Gaver et al. 2004).

address a particular problem, in my case epistemic action, but b) processually underspecified; by this I mean the specific way in which they are to be put to work in a particular situation is intentionally left open for discussion and indeed constitutes an important part of the research and evaluation itself. Some may argue that prototypes inherently hold the same potential, that is to say for users to put them to work differently than intended by the designer, which is obviously impossible to control for. The difference with prototypes, however, is that such other-than-intended tool use is not some unexpected, serendipitous accident, but should be actively designed for by following a minimal reserved approach through suggesting interpretational openness as well as being anticipated and allowed for during field work (see previous chapter). In this regard, they are clearly inspired and informed by a rich body of work¹¹² in the development of cultural probes and, thus, may be argued to be a special type of cultural probe.

Reconceptualising my prototypes as probotypes was also supported by the RTD method, which conceptualises prototypes as more interested in *what might be* (Gaver 2012) rather than designing the *right thing* (Zimmermann et al. 2006). Such an understanding is clearly in line with the commitments and motivations of this research stemming from identifying a variety of issues in medical epistemology that arguably fall within the category of ontological politics as discussed in chapter four. While some suggest that RTD should employ more standardised methods and procedures, including evaluation, with the aim of extensibility and verifiability, such a development would in light of this research be counterproductive (Zimmermann et al. 2006). I speculate that the diagnostic potential of probotypes in the understanding of professional roles and practices would have been less likely to have surfaced. This is because the insights probotypes may produce are better seen as context-specific and, thus, contingent. This, in turn, allows designers to more meaningfully engage with the intricacies of a specialist field (Collins 2014, 73) which may be vital when approaching interaction design from an extended cognitive processes perspective (Hollan et al. 2006, 179). I write vital because this allows, as shown in my case above, to better understand how medical professionals interact with task and work environments, including social structures. In turn, researchers and designers will be able to take into account how professionals with different roles coordinate their interactions and the role tools may play in such an environment. Probotypes may, thus, be defined as a particular diagnostic approach of applying the RTD method in order to elicit

¹¹² The inspiration stems to a large part from the body of work developed by Gaver and colleagues in the Interaction Research Studio at Goldsmiths College, London. See: <http://www.gold.ac.uk/interaction/>. ac. 20.10.2014.

(professional) practices and more precisely diagnose potentials for meaningful interactions within complex processes and environments.

Chapter 8

Discussion of implications and future directions

Brief introduction

In this final chapter I will summarise and discuss the development of this work, its claims and findings, the implications it raises as well as its limitations. The first part will focus on the theoretical review of the literature and the formulation of the research hypothesis. I will especially focus on the issues identified in medical epistemology and evidence-based medicine as well as the implications this has for decision-making in preference-sensitive care situations,¹¹³ such as prostate cancer screening. In turn, the hypothesis and its main claims will be reviewed.

The second part will focus on the development of the practice-based research aspect of the thesis. Here, I will discuss the research method and insights generated during the design phase as well as the research findings with reference to the hypothesis and its main claims. Given the interdisciplinary nature of this project, the aim of this final chapter is to consolidate the main research questions, their methodical and methodological considerations, the findings and the implications it raises.

Research context and hypothesis

This research launched out of the observation that, by 2010, searching for health information online had become the third most “popular” activity among internet users of 18 years and above in the USA (Zickuhr 2010, 3). For many, health information found online affects their medical decision-making. However, as documented (Gigerenzer 2002, 4; Gigerenzer and Gray 2011, 3), there are significant issues with understanding and carefully assessing health information, such as difficulties in understanding statistical information as well as critically assessing the quality of medical information (Kukla 2007, 33). Whether found online or in a decision aid provided by a medical professional, cancer risk information based on which to decide whether or not to conduct screening is probabilistic in nature. As discussed in chapter three, such

¹¹³ To recapitulate: preference-sensitive care relates to medical situations for which there is no single best treatment. However, the implications of the treatment options have significant and diverse effects on patients and their lives; thus, their preferences matter.

information is typically presented as natural frequencies or in a clustered visualisation style privileging magnitudes of risk at the expense of the uncertain character of how it plays out at the individual level. As discussed in chapter four, significant levels of overdiagnosis and overtreatment have been identified in cancer screening (Welch and Black 2010). This means that the diagnostic and preventive treatment interventions help some people by treating all identified (and willing) as potentially at risk, but is also missing others due to false-negative results. Also, in the case of prostate cancer, preventive interventions may entail a variety of iatrogenic effects. These include incontinency, impotence, anxieties as well as bacterial and potentially lethal infections from conducting biopsies (Gale 2011). Unfortunately, evidence-based medicine has witnessed a number of issues, including selective reporting and reporting bias, which leads to a significant distortion of the evidence base (Goldacre 2012; Ioannidis 2005).

Following my analysis, cancer screening decisions are characterised by multifarious complexities. These concern the information environment, such as in the case of a distorted medical evidence base and the probabilistic nature of cancer risk information, which renders such information not only problematic but also raises questions as to its usefulness at the individual level (Politi et al. 2007). While medical decision support as part of shared medical decision-making programmes has become increasingly popular, I have argued in chapter three that it fails to include significant elements relevant to preference-sensitive care situations. As I argue in this thesis, this may be said to be because its design resembles more an information-retrieval model of interaction. With regard to the aims of shared decision-making, and as has also been argued by Kukla (2005, 35), patient autonomy may need to be reconsidered to include a patient's larger healthcare situation as well as personal context. Thus, patients may be better seen as active inquirers and their autonomy supported and facilitated by medical professionals. For Kukla, patient autonomy may be less understood as deriving from pure autonomous learning and knowing in every detail, and more as a kind of collaborative practice.

With respect to patient preferences, Epstein and Peters (2009, 195) argue that, in medical situations characterised by substantial uncertainties and potential "outcomes that have previously not been considered or cannot be imagined," preferences are constructed rather than simply elicited. Within the context of prostate cancer screening and concerning the design of decision aids, I argue that the absence of supporting preference construction and of highlighting the random character of pathology is problematic. This is because prostate cancer screening qualifies as a preference-sensitive care situation

and, thus, there is no single best medical outcome. However, the implications following preventive cancer interventions, such as potential impotency and incontinence, significantly affect a patient's life and should be factored into the decision-making process. Typically, in the absence of previous experience or deliberation, this may prove very difficult, as the patient may simply not have an imagination of what it means to live with such implications. As stated in the introduction and developed in chapter four, the hypothesis for this research, then, was that an exploratory model of interaction is better suited to the context and challenges of redesigning and supporting shared decision-making for prostate cancer screening.

Development of research-through-design work

The design process constitutes the first of two significant elements for discussing the hypothesis. The second element concerns the research findings and implications, which will be discussed in the next part of the chapter.

As I have documented and discussed in chapter six, the design process materialised in a number of prototypical ideas. These were sketched and evaluated for their merit based on an exploratory model of interaction. In contrast to the previously discussed information-retrieval model, an exploratory model of interaction is characterised by iterative, multi-faceted interactions involving incremental learning about the knowledge domain as well as, in some cases, successive clarification of the knowledge goal itself (White 2009). Such a model of interaction may also be said to be more closely aligned with Schön's concept of reflection-in-action (1983). To recapitulate, instead of reflecting on the outcome after the fact, in some situations we reflect "in the midst of action without interrupting it" and, thus, make immediate changes (ibid.).

In the design-process phase of the project I moved from an initially static representation of the false-positive issue in idea #1a stencil and idea #1b grid, to a mapping of an understanding in idea #2a Venn diagram (mapping), to the finally enactable prototype in idea #b Venn diagram (interaction). The exploratory interaction model supported making these steps by guiding the analysis and identification of multiple potentials for epistemic action. Simultaneously, it helped identify the limitations of the prototypical ideas as discussed in detail in chapter six. This successive prototype development, in turn, may also be seen as a reflexive use of the extended mind thesis as part of the research-through-design approach. That is to say, reflection-in-action

throughout the iterative design process was informed by and simultaneously analysed through the evaluative criteria of the prototype's potential for epistemic action. This was achieved by an analysis of the detailed interactions designed and their potential meaningfulness.

Clearly, such an approach is limited by the capacity to anticipate a variety of possible interpretations by other people. Despite this limitation, it seemed productive as, from the design development work, it has also become clear that an information-retrieval model may be seen as a more suitable model of interaction for information re-finding. That is to say, supposing that a person is already familiar with the meaning of a piece of information and, upon re-finding it, can easily identify and act upon it. Thus, the design goal following an information-retrieval approach may be argued to be finding an ideal representation chiefly supportive of one specific interpretation. Clearly, these are not the characteristics that apply to the preference-construction and decision-making context in prostate cancer screening. This is a key point I have argued above and in detail concerning the issues inherent in medical epistemology and evidence-based medicine (chapter two) as well as with regard to the presentation of probabilistic risk information, a reconsideration of the notion of patient autonomy and the need to support the preference-construction process (chapter three).

Following the RTD method, the insights I generated regarding how to redesign shared decision-making (SDM) successively consolidated in what I called a 'minimal reserved approach.' In summary, I characterised such an approach as follows: focusing on the hypothesised core aspects relevant for epistemic action and keeping additional explanatory embedding to a minimum. The latter concerns how the tools will be put to use in a particular context, and will be discussed in the next section. With regard to the former, and in the context of making sense of PSA test results, this means that the enactable visual form should - in principle - allow to simultaneously understand the possibility of a healthy and pathological state when viewed from a population perspective *and* as a binary between being healthy or pathological when viewed from an individual perspective. Furthermore, such an approach recognises the need to elicit an interpretation by medical professionals regarding the multiple possibilities of epistemic action hopefully designed into the tools.

Understanding how the medical professionals would make sense of the tool and put it to use with patients as part of their consultation practice brings me to the discussion regarding explanatory embedding. This was sought by following an approach of underspecifying the design in terms of its explanatory embedding. Underspecification of the design was achieved by multiple

means. Firstly, in the design of the Venn interaction diagram tool, I left a variety of spaces unused. These could be used to add further explanatory or useful notes upon the evaluation of medical professionals or, for example, inviting users to take notes themselves. Secondly, I decided not to include a manual on how to use the tools. Rather, and as part of the evaluation process, I handed the tools to the medical professionals and asked them how they might put the tools to use. Only when asked how I envisioned the tools to be used did I discuss my thoughts on this. Such was the case when, for example, asked by Poyet, the urologist from the Prostate Carcinom Centre whom I had in mind as a potential user of the cards.

By way of this minimal reserved approach, my tools came to act as probe-like prototypes or 'probotypes.' Methodologically, probotypes occupy a middle ground between prototypes and cultural probes. As discussed in chapter seven, they may be best characterised as being semi-directed. While they are functionally directed to address a particular problem just like prototypes, processually they are underspecified, by which I mean the specific way in which they are to be put to work in a particular situation is intentionally left open for discussion and constitutes an important part of the RTD method. Thus, the probotypes uncovered a better understanding of how medical professionals interact with task and work environments, including social structures. In turn, researchers and designers will be able to take into account how professionals with different roles coordinate their interactions and the role tools may play in such an environment.

Probotypes may, thus, tentatively be defined as a particular diagnostic approach of applying the RTD method in order to elicit (professional) practices and more precisely diagnose potentials for meaningful interactions within complex and multi-layered processes and environments. Since evaluating with patients will occur in a later research phase, it will be important and interesting to see in what ways the probotypes might also be useful in that regard. I speculate that they might in fact become productive as a catalyst to uncover opinionated medical professionals (Wennberg 1973). This, in turn, leads me to the discussion of the key findings.

Key findings and their implications

The tools were generally well received and, more importantly, led to interesting insights as noted in chapter seven. One key insight that emerged was the structurally different notion and relevance of time within the decision-making process that was at the heart of the reflection cards I had developed. As noted

by Cornuz and Auer from the University Hospital Lausanne, in contrast to current decision aids, my tools introduced a much longer time-frame for deliberating both the decision itself and the time-frame for which such decisions should be accounted for. This has potentially multiple implications:

Firstly, from the perspective of the screening decision-making person, one immediate implication is viewing the screening decision with regard to such a long-term perspective may entail the need to anticipate a variety of experiences as part of the cancer screening process. As a consequence, this may occasion a shifting of personal tensions and values. In turn, the decision of whether or not to screen may be better viewed as one that may need to be open for reconsideration, hence not necessarily valid indefinitely. The implications this may have will be an interesting research focus for the patient evaluation.

On the one hand, it might be liberating for patients as they could stop with the screening process, should they start to feel uncomfortable with it. On the other hand, I wonder whether people would really 'abandon' the screening process once the spectre of potentially being affected by cancer is eliminated. As commented on by Poyet, the urologist, some people find it difficult to accept that the conducting of a biopsy did not find any cancerous lesions, despite the (false-)positive PSA test results. Furthermore, how people stopping with the cancer screening process will account for it with regard to their social environment will be of interest (Sontag 1977).

Secondly, not everyone may feel comfortable or capable of deliberating such questions solely by himself. Thus, it may be speculated that, without the availability of counseling, such as that for genetic testing, such an approach may be frightening, undesirable or even unacceptable to some (Politi et al. 2007, 688). To some degree, the Cochrane systematic review of decision aids may indeed already have confirmed the discomfort with such deliberation in solitude - given the current decision aids' inability to provide higher satisfaction concerning anxiety when compared to conventional care practices (Cochrane Collaboration 2009, 43).

This aspect may point towards a further potential line of inquiry for the next phase of this research. I wonder whether a digital equivalent of the reflection cards and the notes and comments made by people using them may provide a valuable opportunity to share such deliberations via an online platform. One of the potential positive implications that may follow from such a platform may be that it renders the deliberation and decision process a less lonely one. Furthermore, opening up this process may indeed reveal that it is less linear than it might be thought of or, as I argue, reflected in the design of contemporary decision aids as discussed in chapter three and four. Practically

speaking, the exchange by people involved in or occupied with such screening decision-making on such a platform may show that some people either feel uncomfortable with or even stop the cancer screening process, and document the ways in which they account for it.

Thirdly, from the perspective of medical professionals, two insights stood out. With regard to the reflection cards, all the medical professionals found them useful, but were unable to assess or speculate on the potential reactions that might ensue were they to use them with their patients. As noted in chapter seven, this may point to the absence of practicing preference diagnosis with patients and, importantly, involving patients when doing so. At this point, one may speculate as to whether medical professionals already have the relevant training to support people in the preference-construction process. While some argue that medical professionals “are in general not better placed to make good judgments about moral choices than patients are” (McLeod 2005, 5), Kukla (2007, 32) argues that experienced clinicians may indeed also have moral expertise special to their field of practice. Her argument is that, due to their specialism, they have significant opportunities to reflect on the moral complexities that concern treatment options and medical procedures they conduct on a daily basis. In contrast, and as discussed earlier, many people find themselves confronted with such questions for the first time. It would be my hope that the reflection cards stimulate people considering cancer screening to think about how this might affect them in their personal life and evoke questions and dilemmas they might want to discuss with their trusted GP so as to get their view and expertise on these aspects. Additionally, with regard to the digital platform just discussed, it is imaginable that such expertise may also manifest as part of the dialogue on such a platform. To this end, my design may be seen as hoping to initiate such deliberations and dialogues between people considering cancer screening and their GPs as well as among themselves, offline or online.

The second insight from the perspective of medical professionals concerns the role my prototypes came to serve regarding the ways in which medical professionals with different roles (e.g. a GP or a urologist) see their roles and what is expected of them and their practice. Poyet as a urologist sees many people with an ‘elevated’ PSA test result being referred to him by GPs. His, and presumably that of other urologists, understanding of such referrals is that he is to conduct a prostate biopsy in order to verify whether such elevated levels are in fact due to cancerous lesions or, instead, yet another false-positive PSA test result. Specifically, he thought he could not use the reflection cards in the case of a referral patient, as this would question the medical authority of the referring GP and may negatively affect the trust the

patient has both in his GP as well as medicine more largely. Instead, Poyet suggested that the problem may lie in that PSA tests are conducted without prior thoughtful deliberation with a person concerning the harms and benefits of screening and, thus, preference diagnosis. This could be addressed by setting up a regular consultation hour, which could be attended both by potential prostate screening candidates as well as their GPs before any PSA test is conducted.

One immediate question is why such consultations prior to conducting PSA tests are not yet available. Surprisingly, my suggestion to set up a meeting with Huber, the GP, and Poyet, the urologist, has been met with little interest by Huber yet. Nevertheless, it is noteworthy that Poyet, the urologist, proposed to set up a regular consultation hour prior to PSA testing and suggested that, were preference diagnosis to be conducted, many examinations could be spared. It is somewhat incomprehensible why Huber, would not see such suggestions made by a urologist in a management position of said institution as an opportunity to advance a shared matter of concern. Viewed from a distance, this seems a clear reminder of the strong limitations the tools I have developed may have. Unless there were to be a shared acceptance by medical professionals and an interest in reconsidering current clinical practice of launching cancer-screening processes without prior preference diagnosis, my tools may have limited impact.

Conclusion, limitations and future research directions

Looking at the evaluation sessions might indicate that there is little immediate feedback on whether extended cognitive processes would be supported by the tools I designed. However, in a less direct manner I infer that the medical professionals share the perceived potential of the extended cognitive approach. This is because: a) at the beginning of the evaluation sessions they confirmed to have understood the cognitive approach chosen for designing the tools and had no questions about it or why I had followed such an approach; b) the interactive elements and suggested epistemic action of the tools were easily explained and found useful; and c) consequently, and most importantly, they could easily see themselves using the tools with their patients in consultations. This was further confirmed when Cornuz and Auer highlighted and commended the fundamental reframing of the decision process and time-frame in my tools, which is an important element in supporting extended cognitive processes.

To clarify, the goal of this research was neither to convince medical

professionals nor patients of the veracity of philosophical arguments about extended cognitive processes. Rather, I raised the question of how we could design with EMT in mind within the context of exploratory health decision-making processes. Having medical professionals recognise a potential in the approach taken and the tools designed by being willing to evaluate them in real consultation processes is a very positive sign for this research.

At the same time, there are also clear limitations of what the cards may ultimately be capable of achieving. As Auer remarked, “the cards will not be able to unbiase the doctor.” This comment touches on the issue of preference-sensitive care situations and the relation between a doctor and a patient. With regard to “unbiasing doctors,” while this is something neither the reflection cards nor the Venn interaction tool aim to achieve, it is to be expected that a medical professional with a set opinion on the benefits of prostate cancer screening would likely be hesitant to use my tools in the consultation in the first place.

Whether patients of such a medical professional would come across my tools by other means and how that would influence their relation with such a medical professional is an interesting and open question. Assuming a relatively low difficulty of understanding the tools I have developed, it seems easy to imagine complementing them with explanatory material for introducing people to them directly rather than via a medical professional. Hoping that my tools might also eventually be endorsed by relevant medical institutions, such as the Swiss Cancer League, this would likely make necessary for said skeptical medical professional to take a stand when confronted by a patient having used the tools. Potentially, such a moment of confrontation might also reveal a bit more clearly the degree to which the practice of defensive medicine¹¹⁴ influences the behaviour and suggestions of medical professionals.

In conclusion, and in reflection on my hypothesis, I argue that an exploratory model is indeed better suited to the purposes of supporting shared decision-making in prostate cancer screening. This is because such a model of interaction helps with identifying, understanding and addressing multiple potentials for epistemic action as demonstrated in the design and evaluation process (chapter six and seven). This is much needed in a complex information and decision environment, as I analysed and argued in detail with regard to medical epistemology (chapter two), shared medical decision-making (chapter three) and the ontological politics underlying prostate cancer screening decisions (chapter four). An exploratory model of interaction

¹¹⁴ The practice of suggesting medical treatments that may not be necessary but reduce the risk of the medical professional being legally charged for not conducting them.

requires us to recognise multifarious trajectories of interaction; in turn, it is more open to analysing multiple ways in which an understanding of complex and multi-layered pieces of information may be supported by interacting, interrogating and learning about them successively.

Throughout the development of the design work as well as when evaluating it with medical professionals, the problem and design space I have chosen to engage with has become clearer and much more concrete. In turn, I have come to realise and appreciate the complexity of the design and decision context within which I have aimed to start developing relevant interaction prototypes. The research, development and evaluation process has personally been a highly challenging and satisfactory learning experience.

Following this, and as discussed in this concluding chapter, there are a number of promising lines of inquiry in which I can and will take this research project further. Firstly, I look forward to continuing to refine the research methodology by employing prototypes as a specific approach to the research-through-design method. I will do this by inventing and developing a variety of tools to support epistemic action and preference-construction processes. Secondly, and as suggested by the medical professionals, I may also start to do this for other types of cancer and affected people, such as breast cancer. Lastly, I am also considering taking the tools to the digital realm so as to be able to inquire into the possibilities this may have for opening up the deliberation and decision process beyond the decision-taking individual. All of this will happen in the next phases of this research project when planning and setting up patient evaluation projects with medical professionals.

Bibliography

- Ablin, R. J. 2010. "The Great Prostate Mistake." *New York Times*, March 9. http://www.nytimes.com/2010/03/10/opinion/10Ablin.html?_r=0 (accessed March 1, 2013).
- Ancker, J. S., Y. Senathirajah, R. Kufaka, and J. B. Starren. 2006. "Design Features of Graphs in Health Risk Communication: A Systematic Review." *Journal of the American Medical Informatics Association* 13 (6): 608-618.
- Anderson, Chr. 2006. *The Long Tail: Why the Future of Business is Selling Less of More*. New York: Hyperion Books.
- Armstrong, D. 1995. "The Rise of Surveillance Medicine." *Sociology of Health and Illness* 17 (3): 393-404.
- Auber, T. C. E. 1839. *Traité de Philosophie Médicale*. Paris.
- Ayman, D., and A. D. Goldshine. 1940. "Blood Pressure Determinations by Patients with Essential Hypertension: I. The Difference between Clinic and Home Readings before Treatment." *The American Journal of the Medical Sciences* 200 (4): 465-474.
- Bach-y-Rita, P., C. Collins, F. A. Saunders, B. White, and L. Scadden. 1969. "Vision Substitution by Tactile Image Projection." *Nature* 221 (5184): 963-964.
- Barry, M. J., A. G. Mulley, F. J. Fowler, and J. W. Wennberg. 1988. "Watchful Waiting vs Immediate Transurethral Resection for Symptomatic Prostatism." *Jama* 259 (20): 3010-3017.
- Bates, M. J. 2002. "Toward an Integrated Model of Information Seeking and Searching." *The New Review of Information Behaviour Research* 3: 1-15.
- Bates, D. W., G. J. Kuperman, S. Wang, T. Gandhi, A. Kittler, L. Volk, C. Spurr, et al. 2003. "Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-based Medicine a Reality." *Journal of the American Medical Informatics Association* 10 (6): 523-530.
- BBC 2010. "Experts Scrap Prostate Screening Proposal." *BBC News Health*, December 6. <http://www.bbc.co.uk/news/health-11930979> (accessed March 1, 2013).
- Belkin, N. 2008. "Some(what) Grand Challenges for Information Retrieval." Paper presented at the European Conference on Information Retrieval, Glasgow, March 31.

Berkson, J. 1942. "Tests of Significance Considered as Evidence." *Journal of the American Statistical Association* 37 (219): 325-335.

Blount, M., M. R. Ebling, J. M. Eklund, A. G. James, C. McGregor, N. Percival, K. P. Smith, and S. Daby 2010. "Real-Time Analysis for Intensive Care." *IEEE Engineering in Medicine and Biology Magazine* March/ April: 110-118.

BMJ Group (no date) Shared Decision Making Deciding whether to have a prostate specific antigen (PSA) test.

Bowker, G. C., and S. L. Star. 1999. *Sorting Things Out*. Cambridge: MIT Press.

Boyd, D., and K. Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15 (5): 662-679.

Brawley, O. W., and T. Gansler 2010. "Introducing the 2010 American Cancer Society Prostate Cancer Screening Guideline." *CA: A Cancer Journal for Clinicians* 60 (2): 68-69.

Broder, A. 2002. "A Taxonomy of Web Search." *ACM Sigir Forum* 36 (2): 3-10.

Brooks, R. 1991. "Intelligence without Representation." *Artificial Intelligence* 47: 139-159.

Broussais, F. J. V. 1832. *Physiological Medicine*. Philadelphia: Carey & Lea.

Bryant, R., R. H. Katz, and E. D. Lazowska. 2008. "Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science and Society." Computing Research Association. http://www.cra.org/ccc/docs/init/Big_Data.pdf (accessed February 5, 2013).

Buchanan, R. 2001. "Design Research and the New Learning." *Design Issues* 17 (4): 3 -23.

Burn-Murdoch, J. 2012. "Health Survey for England: the Key Data on Obesity." *The Guardian*, December 20. <http://www.theguardian.com/news/datablog/2012/dec/20/health-survey-england-obesity-trends-data> (accessed January 13, 2013).

Bursztajn, H., Feinbloom, R. I., Hamm, R.M., Brodsky, A. (1981) *Medical choices medical chances*. New York: Delacorte Press.

Butler, D. 2013. "When Google Got Flu Wrong." *Nature*, February 13. <http://www.nature.com/news/when-google-got-flu-wrong-1.12413> (accessed February 15, 2013).

- Canguilhem, G. 1994. *A Vital Rationalist - Selected Writings from Georges Canguilhem*. New York: Zone Books.
- Carel, H. 2008. *Illness*. Durham: Acumen Publishing.
- Castells, M. 1996. *The Rise of the Network Society (Vol 1)*. Oxford: Blackwell Publishers.
- Center for Public Policy and the Social Sciences. 2011. *Shared Decision Making in Health Care Delivery*.
http://rockefeller.dartmouth.edu/shop/shared_decisionmaking_final_v10.pdf
(accessed August 15, 2014).
- Chassin, M. R., R. H. Brook, R. E. Park, J. Keeseey, A. Fink, J. Kosecoff, K. Kahn, N. Merrick and D.H. Solomon. 1986. "Variations in the Use of Medical and Surgical Services by the Medicare Population." *The New England Journal of Medicine* 314: 285-290.
- Christakis, N. A., and J. H. Fowler. 2007. "The Spread of Obesity in a Large Social Network over 32 Years." *New England Journal of Medicine* 357 (4): 370-379.
<http://www.nejm.org/doi/full/10.1056/nejmsa066082> (accessed May 10, 2012).
- Clark, A., and D. Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7-19.
- Clark, A. 1998. *Being There: Putting Brain, Body & World Together Again*. Massachusetts: MIT Press.
- Clark, A. 2003. *Natural Born Cyborgs*. Oxford: Oxford University Press.
- Clark, A. 2007. "Curing Cognitive Hiccups: A Defense of the Extended Mind." *The Journal of Philosophy* 104 (4): 163-192.
- Clark, A. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Cambridge: MIT Press.
- Clark, A. 2009. "Dispersed Selves." *Leonardo Electronic Almanac* 16: 4-5.
- Cochrane, A. L. 1972. *Effectiveness and Efficiency: Random Reflections on Health Services*. London: Royal Society of Medicine Press.
- Cochrane Collaboration. 2014 (last updated, May 26). "Evidence-based health care and systematic reviews."
<http://www.cochrane.org/about-us/evidence-based-health-care>
(accessed February 1, 2014).

Cochrane Collaboration. 2009. *Decision Aids for People Facing Health Treatment or Screening Decisions*. The Cochrane Library: Wiley.
<http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD001431.pub4/abstract>
(accessed February 10, 2014).

Cohen, J. 1994. "The Earth is Round ($p < .05$)." *American Psychologist* 49 (12): 997-1003.

Collins, H. 2014. *Are We All Scientific Experts Now?* Malden: Polity Press.

Cook, A. R., M. I. Chen, and R. T. P. Lin. 2010. "Internet Search Limitations and Pandemic Influenza, Singapore." *Emerging Infectious Diseases* 16 (10): 1647.
http://wwwnc.cdc.gov/eid/article/16/10/10-0840_article.htm (accessed April 15, 2012).

Cornuz, J., B. Kuenzi, and T. Krones, 2011. "Shared Decision Making Development in Switzerland: Room for Improvement!" *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 105 (4): 296 - 299.

Dartmouth Atlas Project Topic Brief. 2007 (January 15). "Preference-Sensitive Care."
http://www.dartmouthatlas.org/downloads/reports/preference_sensitive.pdf (accessed March 10, 2013).

Diebold, F. 2003. "'Big Data' Dynamic Factor Models for Macroeconomic Measurement and Forecasting." In *Advances in Economics and Econometrics*, edited by M. Dewatripont, L.P. Hansen, and S. Turnovsky, 115-122. Cambridge: Cambridge University Press.

Diebold, F. 2012. "On the Origin(s) and Development of the Term, Big Data." In *PIER Working Paper* 12-037.
<http://economics.sas.upenn.edu/sites/economics.sas.upenn.edu/files/12-037.pdf>
(accessed March 15, 2013).

Djulgovic, M., R. J. Beyth, M. M. Neuberger, T. L. Stoffs, J. Vieweg, B. Djulgovic, and P. Dahm. 2010. "Screening for Prostate Cancer: Systematic Review and Meta-analysis of Randomised Controlled Trials." *British Medical Journal* 341: c4543.

Doll, R., and A. B. Hill. 1950. "Smoking and Carcinoma of the Lung." *British Medical Journal* 2 (4682): 739-748.

Doll, R., and A. B. Hill. 1954. "The Mortality of Doctors in Relation to their Smoking Habits." *British Medical Journal* 1 (4877): 1451-1455.

Donald, M. 1991. *Origins of the Modern Mind*. Harvard: Harvard University Press.

Dourish, P. 2004. *Where the Action is: The Foundations of Embodied Interaction*. Massachusetts: MIT Press.

Dreyfus, H. 1979. *What Computers Can't Do*. New York: Harper-Collophon.

Dugas, A. F., Y. H. Hsieh, S. R. Levin, J. M. Pines, D. P. Mareiniss, A. Mohareb, Chr. Gaydos, T. Perl and R. Rothman. 2012. "Google Flu Trends: Correlation with Emergency Department Influenza Rates and Crowding Metrics." *Clinical Infectious Diseases* 54 (4): 463-469.

<http://cid.oxfordjournals.org/content/early/2012/01/02/cid.cir883.short> (accessed April 15, 2012).

Dwan, K., D. G. Altman, J. A. Arnaiz, J. Bloom, A. W. Chan, E. Cronin, E. Decullier, et al. 2008. "Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias." *PLoS One* 3 (8): e3081.

Eddy, D.M. 1982. "Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities." In *Judgment under Uncertainty: Heuristics and Biases*, edited by D. Kahneman, P. Slovic, and A. Tversky, 249 - 267. New York: Cambridge University Press.

Eddy, D. M. 1984. "Variations in Physician Practice: The Role of Uncertainty." *Health Affairs* 3 (2): 74-89.

Eddy, D. M. 1990. "Practice Policies: Where do They Come From?" *JAMA: The Journal of the American Medical Association* 263 (9): 1265-1275.

Eddy, D. M. 2005. "Evidence-based Medicine: A Unified Approach." *Health Affairs* 24 (1): 9-17.

Edwards, D., C. Baum, and N. Morrow-Howell. 1994. "Home Environments of Inner City Elderly with Dementia: Do They Facilitate or Inhibit Function?" *Gerontologist* 34 (1): 1-64.

Epstein, A.M. 1990. "The Outcomes Movement - Will it Get Us Where We Want to Go?" *The New England Journal of Medicine* 323 (4): 266-270.

Epstein, R. M., and E. Peters. 2009. "Beyond Information: Exploring Patients' Preferences." *Journal of the American Medical Association* 302 (2): 195-197.

Epstein, R. M., and R. L. Street. 2011. "The Values and Value of Patient-Centred Care." *Annals of Family Medicine* 9 (2): 100-103.

Esserman, L. J. and I. M. Thompson. 2010. "Solving the Overdiagnosis Dilemma."

Journal of the National Cancer Institute 102 (9): 582-583.

Esserman, L. J., I. M. Thompson, and B. Reid. 2013. "Overdiagnosis and Overtreatment in Cancer: An Opportunity for Improvement." *JAMA* 310 (8): 797-798.

Esposito, E. 2007. *Die Fiktion der wahrscheinlichen Realität*. Frankfurt: Suhrkamp.

Eyler, J. M. 1979. *Victorian Social Medicine - The Ideas and Methods of William Farr*. London: John Hopkins University Press.

Farr, W. 1990. "On a Method of Determining the Danger and the Duration of Diseases at Every Period of Their Progress." In *The Taming of Chance*, edited by I. Hacking, 72 - 79. Cambridge: University Press.

Feuz, M., M. Fuller, and F. Stalder. 2011. "Personal Web Searching in the Age of Semantic Capitalism: Diagnosing the Mechanisms of Personalization." *First Monday* 16 (2).
<http://ojs-prod-lib.cc.uic.edu/ojs/index.php/fm/article/view/3344/2766> (accessed April 21, 2014).

Fisher, R. 1925. *Statistical Methods for Research Workers*. Ann Arbor: University of Michigan Press.

Forlizzi, J., C. DiSalvo, J. Bardzell, I. Koskinen, and S. Wensveen. 2011. "Quality Control: A Panel on the Critique and Criticism of Design Research." CHI 2011 Panel, 823-826.

Foucault, M. 1973. *The Birth of the Clinic*. London: Tavistock Publications.

Fox S., 2006. "Online Health Search 2006." Pew Internet & American Life Project.
http://www.pewinternet.org/files/old-media/Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf (accessed April 20, 2014).

National Heart, Lung and Blood Institute. *Framingham Heart Study*.
<http://www.framinghamheartstudy.org/> (accessed February 1, 2013).

Frayling, Chr. 1993. *Research in Art and Design*. Royal College of Art Research Papers 1 (1): 1-5.

Freedman, D.H. 2010. "Lies, Damned Lies, and Medical Science." *The Atlantic*, Nov.
<http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/8269/> (accessed May 2, 2012).

Fuller, M. 2003. *Behind the Blip*. New York: Autonomedia.

Fuller, M. 2008. *Software Studies – A Lexicon*. Cambridge: MIT Press.

Fuller, M., and S. Matos. 2011. "Feral Computing, from Ubiquitous Calculation to Wild Interactions." *Fibreculture Journal* 19.

<http://nineteen.fibreculturejournal.org/fcj-135-feral-computing-from-ubiquitous-calculation-to-wild-interactions/> (accessed February 5, 2012).

Fuller, M. and Goffey Andrew. 2012. *Evil Media*. Cambridge: MIT Press.

Gale, J. 2011. "Prostate Biopsy Can Be Risky: More Men Getting Ill from Drug-resistant Infections." *The Post and Courier*, May 9.

<http://www.postandcourier.com/article/20110509/PC1602/305099972> (accessed May 2, 2013).

Garfinkel, H. and Sacks, Harvey. (1970) *On formal structures of practical actions*. In J.McKinney and E.Tyrakian (eds.), *Theoretical sociology: Perspectives and development* (pp.337-366). New York: Appleton Century Crofts.

Gaver, W., T. Dunne, and E. Pacenti. 1999. "Cultural Probes." *Interactions* 6 (1): 21-29.

Gaver, W., A. Boucher, S. Pennington, and B. Walker. 2004. "Cultural Probes and the Value of Uncertainty." *Interactions* 11 (5): 53-56.

Gaver, W. 2011. "Making Spaces: How Design Workbooks Work." In *Proceedings of the Special Interest Group in Computer-Human Interaction (SIGCHI) Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, 1551-1560.

Gaver, W. 2012. "What Should We Expect from Research through Design?" In *Proceedings of the Special Interest Group in Computer-Human Interaction (SIGCHI)*, 937-946.

Gigerenzer, G. 2002a. *Reckoning with Risk: Learning to Live With Uncertainty*. London: Allen Lane.

Gigerenzer, G. 2002. *Calculated Risks: How to Know When Numbers Deceive You*. New York: Simon & Schuster.

Gigerenzer, G., and U. Hoffrage. 1995. "How to Improve Bayesian Reasoning Without Instruction: Frequency Formats." *Psychological Review* 102 (4): 684-704.

Gigerenzer, G., W. Gaissmaier, E. Kurz-Milcke, L. M. Schwartz, and S. Woloshin. 2007. "Helping Doctors and Patients Make Sense of Health Statistics." *Psychological Science in the Public Interest* 8 (2): 53-96.

Gigerenzer, G., Krauss, S. and Vitouch, O. 2004. "The Null Ritual." In *The Sage Handbook of Quantitative Methodology for the Social Sciences*, edited by D. Kaplan, 391-408. Sage Publications.

Gigerenzer, G., and M. J. A. Gray. 2011. "Launching the Century of the Patient." In *Better Doctors, Better Patients, Better Decisions: Envisioning Health Care 2020*, edited by Gerd Gigerenzer and J.A. Muir Gray, 3-28. Cambridge, MA: MIT Press.

Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant. 2008. "Detecting Influenza Epidemics Using Search Engine Query Data." *Nature* 457 (7232): 1012-1014.

Goel, S., J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts. 2010. "What Can Search Predict." *WWW 2010*, April 26-30, 2010, Raleigh, North Carolina.

Goldacre, B. 2012. *Bad Pharma - How Drug Companies Mislead Doctors and Harm Patients*. London: Fourth Estate Harper Collins.

Goldacre, B. 2012. "The Information Architecture of Medicine is Broken." Strata Talk. O'Reilly Media
http://www.youtube.com/watch?v=AK_EUKJyusg (accessed February 1, 2013).

Goode, E. 2011. "Sending the Police before There's a Crime." *New York Times*, August 15.
http://www.nytimes.com/2011/08/16/us/16police.html?_r=1& (accessed February 10, 2012).

Goodman, St. N. 1999a. "Toward Evidence-based Medical Statistics. 1: The P Value Fallacy." *Annals of Internal Medicine* 130 (12): 995-1004.

Goodman, St. N. 1999b. "Toward Evidence-based Medical Statistics. 2: The Bayes Factor." *Annals of Internal Medicine* 130 (12): 1005-1013.

Graham, M. 2002. *Democracy by Disclosure - The Rise of Technopopulism*. Washington: Brookings Institution.

Graham, St., and S. Marvin. 2001. *Splintering Urbanism: Networked Infrastructures, Technological Mobilities and the Urban Condition*. London: Routledge.

Gray, M. 1997. *Evidence-Based Healthcare and Public Health*. Edinburgh: Churchill Livingstone Elsevier.

Greene, J. A. 2007. *Prescribing by Numbers: Drugs and the Definition of Disease*. Baltimore: John Hopkins University Press.

Greenhalgh, T. 2010. *How to Read a Paper: The Basics of Evidence-based Medicine*, 5th ed. Oxford: Blackwell.

Greenhalgh, T., J. Howick, and N. Maskrey. 2014. "Evidence Based Medicine: A Movement in Crisis?" *British Medical Journal* 348: 3725.

Griffiths, P. E., and K. Stotz. 2000. "How the Mind Grows: A Developmental Perspective on the Biology of Cognition" *Synthese* 122 (1-2): 29-51.

Hacking, I. 1990. *The Taming of Chance*. Cambridge: Cambridge University Press.

Hacking, I. 2006. "Making up People." *London Review of Books* 28 (16): 23-26.

Hammond, J. L., and B. Hammond. 1917. *The Town Labourer 1760-1832: The New Civilisation*. London: Longmans, Green and Co.

Harford, T. 2013. "Why We Do What We Do." *Financial Times*, January 28.
<http://www.ft.com/cms/s/2/76e593a6-28eb-11e0-aa18-00144feab49a.html#axzz1lAG3KbL5> (accessed April 7, 2013).

Harrison, S., D. Tatar, and Ph. Sengers. 2007. "The Three Paradigms of HCI." Virginia Tech.

Havalais, A. 2009. *Search Engine Society*. Cambridge: Polity Press.

Hayles, K. 2012. *How We Think: Digital Media and Contemporary Technogenesis*. Chicago: University of Chicago Press.

Health Dialog, PSA test decision aid, [no date].
<http://www.healthdialog.com/psa> (accessed March 12, 2013).

Hearst, M. 2009. *Search User Interfaces*. Cambridge: Cambridge University Press.

Heidegger, M. 2010. *Being and Time: A Translation of Sein und Zeit*. Translated by J. Stambaugh. State University of New York

Heijnsdijk E. A. M, E. M. Wever, and A. Auvinen. 2012. "Quality-of-life Effects of Prostate-Specific Antigen Screening." *New England Journal of Medicine* 367: 595-605.

Hollan, J., E. Hutchins, and M. Kirsh. 2000. "Distributed Cognition: Toward a New Foundation for Human-computer Interaction Research." *ACM Transactions of*

Computer-Human Interaction 7 (2): 174-196.

Holmes, B. 2007. "Future Map - Or: How the Cyborgs Learned To Stop Worrying and Love Surveillance." *Deriva Continental*.

<http://brianholmes.wordpress.com/2007/09/09/future-map/> (accessed April 15, 2012).

Honjo, K. 2004. "Social Epidemiology: Definition, History, and Research Examples." *Environmental Health and Preventive Medicine* 9: 193-199.

Huber, F., A. Weber, and Chr. Marti. 2012. *Prostatavergrößerung. Symptome, Abklärungen, Behandlung*. Zürich: Medix.

Hutchins, E. 1995a. *Cognition in the Wild*. Massachusetts: MIT Press.

Hutchins, E. 1995b. "How a Cockpit Remembers its Speeds." *Cognitive Science* 19 (3): 265-288.

Ingwersen, P., and K. Järvelin. 2005. *The Turn: Integration of Information Seeking and Retrieval in Context (The Information Retrieval Series)*. Springer-Verlag New York.

Innis, H. 1948. *Empire and Communications*. Toronto: Dundurn Press.

Introna, L., and H. Nissenbaum. 2000. "Shaping the Web: Why the Politics of Search Engines Matters." *Information Society* 16 (3): 169-185.

<http://dx.doi.org/10.1080/01972240050133634> (accessed February 20, 2012).

Ioannidis, J. P. 2005 "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research." *JAMA: the Journal of the American Medical Association* 294 (2): 218-228.

<http://jama.jamanetwork.com/article.aspx?articleid=201218> (accessed April 10, 2012).

Jadad, A.R., and M. W. Enkin. 2007. *Randomised Controlled Trials: Questions, Answers and Musings*. Oxford: Blackwell Publishing.

Jansen, B. J., A. Spink, and T. Saracevic. 2000. "Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web." *Information Processing and Management* 36 (2000): 207-227.

Jeremijenko, N., and B. Bratton. 2008. "Suspicious Images, Latent Interfaces." *The Situated Technology Pamphlet Series*, 3.

<http://situatedtechnologies.net/> (accessed September 10, 2012).

Jones, C., and R. Porter. 2001. *Reassessing Foucault: Power, Medicine and the Body*, 2nd ed. London: Routledge.

Kahneman, D., and A. Tversky. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science New Series* 185 (4157): 1124-1131.

Kelley, D., and J. Rinn. 2012. "Transposable Elements Reveal a Stem Cell-specific Class of Long Noncoding RNAs." *Genome Biology* 13 (11): R107.

Kirsh, D., and P. Maglio. 1994. "On Distinguishing Epistemic from Pragmatic Action." *Cognitive Science* 18(4): 513-549.

Keys, A., F. Fidanza, M. J. Karvonen, N. Kimura, and H. L. Taylor. 1972. "Indices of Relative Weight and Obesity." *Journal of Chronic Diseases* 25 (6): 329-343.

Kobielus, J. 2012. "Big Data Analytics Helps Researchers Drill Deeper into Multiple Sclerosis." *IBM-The Big Data Hub*, May 14.

<http://www.ibmbigdatahub.com/blog/big-data-analytics-helps-researchers-drill-deeper-multiple-sclerosis> (accessed: February 1, 2014).

Krieger, N. 2001. "Historical Roots of Social Epidemiology: Socioeconomic Gradients in Health and Contextual Analysis." *International Journal of Epidemiology* 30 (4): 899-900.

Kuhlthau, C. C. 1991. "Inside the Search Process: Information Seeking from the User's Perspective." *Journal of the American Society for Information Science* 42: 361-371.

Kuhlthau, C. C. 2004. *Seeking Meaning: A Process Approach to Library and Information Services*, 2nd ed. Westport, CT: Libraries Unlimited.

Kukla, R. 2005. "Conscientious Autonomy: Displacing Decisions in Health Care." *Hastings Center Report* 35 (2): 34-44.

Kukla, R. 2007. "How Do Patients Know?" *Hastings Center Report* 37 (5): 27-35.

Lambert, R. 1963. *Sir John Simon and English Social Administration*. Bristol: Macgibbon & Kee.

Lane, J. A., F. C. Hamdy, R. M. Martin, E. L. Turner, D. E. Neal, and J. L. Donovan. 2010. "Latest Results from the UK Trials Evaluating Prostate Cancer Screening and Treatment: The CAP and ProtecT Studies." *European Journal of Cancer* 46 (17): 3095-3101.

Langville, A.N., and C. D. Meyer. 2006. *Google's PageRank and Beyond*. Princeton: Princeton University Press.

Latour, B. 2007. "Beware, Your Imagination Leaves Digital Traces."

<http://www.bruno-latour.fr/sites/default/files/P-129-THES-GB.pdf> (accessed July 15, 2012).

Law, J. (1997) *Aircraft Stories: Decentering the Object in Technoscience*. Keele University, Keele.

Law, J., and M. Urry. 2005. "Enacting the Social." *Economy and Society* 33 (3): 390-410.

Le Fanu, J. 2011. "Statins for All - and Billions for Drug Firms." *The Telegraph*, January 19.

<http://www.telegraph.co.uk/health/8270156/Statins-for-all-and-billions-for-drug-firms.html> (accessed April 20, 2012).

Lehrer, J. 2010. "The Truth Wears off." *The New Yorker*, December 13.

<http://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off> (accessed February 10, 2014).

Lipkus, I. M. 2007. "Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations." *Medical Decision Making* Sept/Oct: 696-713.

Lipkus, I. M., and J. G. Hollands. 1999. "The Visual Communication of Risk." *Journal of The National Cancer Institute Monographs* 25: 149-163.

Lochlan, J.S. 2013. *Malignant: How Cancer Becomes Us*. Berkley: University of California Press.

Lodor, J., L. Bunt, and J.C. Wyatt. 2013. *Doctor Know: A Knowledge Commons in Health*. Nesta.

http://www.nesta.org.uk/publications/assets/features/doctor_know_a_knowledge_commons_in_health (accessed March 20, 2013).

Lumen N., V. Fonteyne, and G. De Meerleert. 2012. "Population Screening for Prostate Cancer: An Overview of Available Studies and Meta-analysis." *International Journal of Urology* 19: 100-108.

Lyon, D. 2002. "Everyday Surveillance: Personal Data and Social Classifications" *Information, Communication and Society* 5: 242-257.

Lyon, D. 2007. *Surveillance Studies: An Overview*. Cambridge: Polity.

Machlup, F.1972. *The Production and Distribution of Knowledge in the United States*. Princeton: Princeton University Press.

Manovich, L. 2008. "Cultural Analytics: Analysis and Visualisation of Large Cultural Data Sets." www.manovich.net (accessed March 11, 2011).

Manovich, L. 2011. "Trending: the Promises and the Challenges of Big Social Data." In *Debates in the Digital Humanities*.
http://www.manovich.net/DOCS/Manovich_trending_paper.pdf (accessed March 15, 2012).

Marchionini, G., 2006. "Exploratory Search: From Finding to Understanding." *Communications of the ACM* 49 (4): 41-46.

Maturana, H., and F. Varela, F. 1987. *The Tree of Knowledge*. Boston, Shambhala.

McLeod, C. reply to Kukla, letters section, Hastings Center Report 35, no. 5 (2005): 5.

McPherson, K., J. E. Wennberg, O. B. Hovind, and P. Clifford. 1982. "Small-area Variations in the Use of Common Surgical Procedures: An International Comparison of New England, England, and Norway." *The New England Journal of Medicine* 307 (21): 1310.

Miller, D., and D. Slater. 2001. *The Internet: An Ethnographic Approach*. London: Bloomsbury Academic.

Milton, G. W. 1973. "Self-willed Death or the Bone-pointing Syndrome." *The Lancet*, June: 1435-1436.
<http://www.ncbi.nlm.nih.gov/pubmed/4122997> (accessed May 1, 2013).

Mol, A. 1998. "Missing Links, Making Links: the Performance of Some Atherosclerosis." In *Differences in Medicine*, edited by M. Berg, and A. Mol, 144-165. Durham: Duke University Press.

Mol, A. 1999. "Ontological Politics: A Word and Some Questions." *The Sociological Review* 47 (S1): 74-89.

Mol, A. 2008. *The Logic of Care*. London: Routledge.

Mol, A., and J. Law. 2004. "Embodied Action, Enacted Bodies: The Example of Hypoglycaemia." *Body & Society* 10: 43-62.

Moore, G.E. 1965. "Cramming More Components onto Integrated Circuits." *Electronics Magazine*, 4: 114 – 117.

Morville, P., and J. Callender. 2010. *Search Patterns: Design for Discovery*.

Sebastopol: O'Reilly Media.

Moynihan, R., J. Doust, and D. Henry. 2012. "Preventing Overdiagnosis: How to Stop Harming the Healthy." *British Medical Journal* 344: e3502.

Muir Gray, J. A. 1997. *Evidence-based Healthcare: How to Make Health Policy and Management Decisions*. London: Churchill Livingstone.

Mulley, A., Chr. Trimble, and G. Elwyn. 2012. "Stop the Silent Misdiagnosis: Patients' Preferences Matter." *British Medical Journal* 345: e6572.

National Science Foundation. 2007. *Cyberinfrastructure Vision for 21st Century Discovery*. March.
<http://www.nsf.gov/pubs/2007/nsf0728/> (accessed February 10, 2012).

Noë, A. 2009. *Out of Our Heads*. New York: Hill and Wang.

O'Connor, A. M., C. L. Bennett, D. Stacey, M. Barrey, N. F. Col, V. A. Eden, R. Entwistle, et al. 2009. "Decision Aids for People Facing Health Treatment or Screening Decisions." *The Cochrane Collaboration*, John Wiley & Sons.

Olin L. S., and L. Sachs. 2001. "Normality, Risk and the Future: Implicit Communication of Threat in Health Surveillance." *Sociology of Health & Illness* 23 (4): 497-516.

Ortiz, J. R., H. Zhou, D. K. Shay, K. M. Neuzil, A. L. Fowlkes, and C. H. Goss. 2011. "Monitoring Influenza Activity in the United States: A Comparison of Traditional Surveillance Systems with Google Flu Trends." *PLoS One* 6 (4): e18687.
<http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0018687>
(accessed October 15, 2011).

Page, L., and S. Brin. 1998. "The Anatomy of a Large-scale Hypertextual Web Search Engine." *Computer Networks and ISDN Systems* 30 (1-7): 107-117.

Pan, B., H. Hembrooke, Th. Joachims, L. Lorigo, G. Gay, L. and Granka L. 2007. "In Google We Trust: Users' Decisions on Rank, Position, and Relevance." *Journal of Computer Mediated Communication* 12 (3): 801-823.

Pinker, St. 2003. *The Blank Slate*. London: Penguin Books.

Pittman, D. 2012. "Big Data in Healthcare: Real-time Health Monitoring and Intervention." IBM.
<http://de.slideshare.net/davidpittman1/big-data-in-healthcare-realtme-health-monitoring-and-intervention> (accessed October 10, 2012).

- Polanyi, M. 1958. *Personal Knowledge: Towards a Post-critical Philosophy*. Chicago: Routledge & Kegan Paul.
- Politi, M. C., P. K. J. Han, and N. F. Col. 2007. "Communicating the Uncertainty of Harms and Benefits of Medical Interventions." *Medical Decision Making* 27: 681-695.
- Porter, R. 1997. *The Greatest Benefit to Mankind: A Medical History of Humanity*. New York: Cambridge University Press.
- Porter, Th. M. 1988. *The Rise of Statistical Thinking, 1820-1900*. Princeton: Princeton University Press.
- Porter, R. 2006. *The Cambridge History of Medicine*. New York: Cambridge University Press.
- Pratt, C. M., and L. A. Moyer. 1990. "The Cardiac Arrhythmia Suppression Trial: Background, Interim Results and Implications." *The American Journal of Cardiology* 65 (4): 20-29. <http://circ.ahajournals.org/content/91/1/245.short> (accessed May 2, 2012).
- Pschyrembel, W. 2014. *Pschyrembel Klinisches Wörterbuch*. Walter de Gruyter.
- Quetelet, A. 1842. *A Treatise on Man and the Development of His Faculties*. London: W.S. Orr and Company.
- Rockhill, B. 2005. "Theorizing About Causes at the Individual Level While Estimating Effects at the Population Level." *Epidemiology* 16 (1): 124-129.
- Rogers, R. 2009. "The Googlization Question, and the Inculpable Engine." http://www.govcom.org/publications/full_list/rogers_inculpable_engine_2009.pdf (accessed April 21, 2014).
- Romero-Corral, A., V. M. Montori, V. K. Somers, J. Korinek, R. J. Thomas, Th. G. Allison, F. Mookadam, and F. Lopez-Jimenez. 2006. "Association of Bodyweight with Total Mortality and with Cardiovascular Events in Coronary Artery Disease: A Systematic Review of Cohort Studies." *The Lancet* 368 (9536): 666-678. <http://www.thelancet.com/journals/lancet/article/PIIS0140673606692519/abstract> (accessed February 20, 2013).
- Rose, N. 1998. "Life, Reason and History: Reading Georges Canguilhem Today." *Economy and Society* 27 (2-3): 154-170.
- Rose, N. 2006. *Politics of Life Itself*. Woodstock: Princeton University Press.

- Rose, N. 2007. "Beyond Medicalisation." *The Lancet* 369 (9562): 700-702.
- Rose, D. E., and D. Levinson. 2004. "Understanding User Goals in Web Search." *Proceedings on the 13th international Conference on World Wide Web*, 13-19.
- Rosen, G. 1974. *From Medical Police to Social Medicine*. New York: Science History Publications.
- Rothman, A., and M. T. Kiviniemi. 1999. "Treating People With Information: An Analysis and Review of Approaches to Communicating Health Risk Information." *Journal of the National Cancer Institute Monographs* 25: 44-51.
- Rothwell, P.M. 2005. "External Validity of Randomised Controlled Trials: 'To Whom Do the Results of this Trial Apply?'" *The Lancet* 365 (9453):82-93.
<http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2804%2917670-8/abstract> (accessed May 2, 2012).
- Rowan, J. 2011. "Transforming Behaviour Change: Beyond Nudge and Neuromania." London: Royal Society Supporting the Arts, Nov.
- Sackett, D. L., W. M. Rosenberg, J. A. Gray, R. B. Haynes, and W. S. Richardson. 1996. "Evidence-based medicine: What it is and what it isn't." *British Medical Journal*, 312 (7023): 71-72.
- Sandblom, G., E. Varenhorst, J. Rosell, O. Löfman, and P. Carlsson. 2011. "Randomised Prostate Cancer Screening Trial: 20 Year Follow-up." *British Medical Journal* 342: d1539.
<http://www.bmj.com/content/342/bmj.d1539> (accessed March 1, 2013).
- Satz, D., and J. Ferejohn. 1994. "Rational Choice and Social Theory." *Journal of Philosophy* 9102: 71-87.
- Schmidt, F. L. 1996. "Statistical Significance Testing and Cumulative Knowledge in Psychology: Implications for Training of Researchers." *Psychological Methods* 1 (2):115-129.
- Schmidt, F. L., and J. E. Hunter 1997. "Eight Common but False Objections to the Discontinuation of Significance Testing in the Analysis of Research Data." *In What if There Were No Significance Tests*, Lisa A. Harlow, Stanley A. Mulaik, and James H. Steiger (Eds.), 37-64. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schön, D. 1983. *Educating the Reflective Practitioner*. Basic Books.
- Sengers, P., and W. Gaver, W. "Staying Open to Interpretation: Engaging Multiple

Meanings in Design and Evaluation.” Paper presented at the Conference of Computer Human Interactions, San Jose, April, 2007.

Shirky, C. 2005. “Ontology is Overrated: Categories, Links, and Tags.” Clay Shirky's Writings about the Internet.

http://www.shirky.com/writings/ontology_overrated.html?goback=.gde_1838701_member_179729766 (accessed February 20, 2014).

Shryock, R.H. 1936. *The Development of Modern Medicine*. London: The University of Wisconsin Press.

Shryock, R. H. 1961. “The History of Quantification in Medical Science.” *Isis* 52 (2): 215-237.

Smart, P., P. C. Engelbrecht, D. Braines, M. Strub, and Ch. Giammanco. 2010. “The Network-Extended Mind.” In *Network Science for Military Coalition Operations: Information Extraction and Interaction*, edited by Dinesh Verma, 191-236. Hershey, Pennsylvania: IGI Global.

Smart, P. 2012. “The Web-Extended Mind.” *Metaphilosophy* 43 (4): 446-463.

Smith, B., W. Ceusters, B. Klagges, J. Köhler, A. Kumar, J. Lomax, Chr. Mungal, et al. 2005. “Relations in Biomedical Ontologies.” *Genome Biology* 6: R46. <http://genomebiology.com/content/6/5/R46> (accessed March 25, 2014).

Sontag, S. 1977. *Illness as Metaphor*. London: Penguin Books.

Spiegelhalter, D., M. Pearson, and I. Short. 2011. “Visualizing Uncertainty about the Future.” *Science* 333: 1393-1400.

Stalder, F., and C. Mayer. 2009. “The Second Index: Search Engines, Personalization and Surveillance.” In *Deep Search: The Politics of Search Beyond Google*, edited by K. Becker, and F. Stalder. Innsbruck: Studien Verlag.

Sterling, T. D. 1959. “Publication Decisions and their Possible Effect on Inferences Drawn from Tests of Significance - or Vice Versa.” *The American Statistician* 54 (285): 30-34.

Sterling, T. D., W. L. Rosenbaum, and J. J. Weinkam. 1995. “Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa.” *The American Statistician* 49 (1): 108-112.

Suchman, L. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge: Cambridge University Press.

Susser, M., and Z. Stein. 2009. *Eras in Epidemiology - The Evolution of Ideas*. Oxford: Oxford University Press.

Swiss Medical Board. 2014. "Swiss Medical Board stellt systematisches Mammographie Screening in Frage." (February 2, 2014). http://www.medicalboard.ch/fileadmin/docs/public/mb/medienmitteilungen/2014-02-02_Medienmitteilung_Bericht_Mammographie-Screening_def.pdf (accessed March 1, 2014).

Tang, D., A. Agarwal, D. O'Brien, and M. Meyer. 2010. "Overlapping Experiment Infrastructure: More, Better, Faster Experimentation." In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 17-26. ACM.

Thaler, R.H., and C. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth and Happiness*. London: Penguin Books.

The Economist. 2012. "The Origin of Species?" December, 1. <http://econ.st/U56g2H> (accessed February 13, 2013).

Timmermans, M., and M. Berg, M. 2003. *The Gold Standard: The Challenge of Evidence-based Medicine and Standardization in Health Care*. Philadelphia: Temple University Press.

Topol, E. 2012. *The Creative Destruction of Medicine*. New York: Basic Books.

Touraine, A. 1974. *The Post-Industrial Society: Tomorrow's Social History: Classes, Conflicts and Culture in the Programmed Society*. London: Willwood House.

Tukey, J. W. 1980. "We Need Both Exploratory and Confirmatory." *The American Statistician* 34 (1): 23-25.

U.S. Preventive Services Task Force. 2012. "Infographic: Benefits and Harms of PSA Screening for Prostate Cancer." NCI Cancer Bulletin 9 (23). <http://www.cancer.gov/ncicancerbulletin/112712/page12> (accessed January 13, 2014).

Wallsten, T., D. Budescu, A. Rapoport, R. Zwick, and B. Forstyth. 1986. "Measuring the Vague Meanings of Probability Terms." *Journal of Experimental Psychology* 115 (4): 348-65.

Welch, H. G., and C. W. Black. 2010. "Overdiagnosis in Cancer." *Journal of the National Cancer Institute* 102 (9): 605-613.

Weiss, Sh. M., and N. Indurkha. 1998. *Predictive Data Mining: A Practical Guide*. San Francisco: Morgan Kaufmann Publishers.

Wenger, E. 1999. *Communities of Practice*. Cambridge: Cambridge University Press.

Wennberg, J. 1973. "Small Area Variations in Health Care Delivery." *Science* 182 (4117): 1102-1108.

Wennberg J. 1982. "Small-Area Variations in the Use of Common Surgical Procedures: An International Comparison of New England, England, and Norway." *The New England Journal of Medicine* 307 (21): 1310-1314.

Wennberg, J. 1984. "Dealing with Medical Practice Variations: A Proposal for Action." *Health Affairs* 3 (2): 6-32.

Wennberg, J. 2002. "Unwarranted Variations in Healthcare Delivery: Implications for Academic Medical Centres." *British Medical Journal* 325: 961-964.

Wheeler, M. 2011. "Thinking beyond the Brain: Educating and Building, from the Standpoint of Extended Cognition." *Computational Culture*, Nov. <http://computationalculture.net/article/beyond-the-brain> (accessed October 15, 2012).

Whelan, T., A. Gafni, C. Charles, and M. Levine. 2000. "Lessons Learned from the Decision Board: A Unique and Evolving Decision Aid." *Health Expectations* 3: 69-76.

White, R. W., and R. A. Roth. 2009. *Exploratory Search: Beyond the Query-Response Paradigm. Synthesis Lectures on Information Concepts, Retrieval, and Services*. Morgan & Claypool Publishers.

White, R. W., St. M. Drucker, G. Marchionini, M. Hearst, and M. C. schraefel. 2007. "Exploratory Search and HCI: Designing and Evaluating Interfaces to Support Exploratory Search Interaction." In Proceedings of SIGCHI.

Wilson, T. D. 1999. "Models in Information Behavior Research." *Journal of Documentation* 55 (3): 249-270.

World Health Organisation Expert Consultation. 2004. "Appropriate Body-mass Index for Asian Populations and its Implications for Policy and Intervention Strategies." *The Lancet* 363 (9403): 157-163. <http://europepmc.org/abstract/MED/14726171> (accessed January 13, 2013).

Žižek, S. 2013. "The Interpassive Subject." <http://www.egs.edu/faculty/slavoj-zizek/articles/the-interpassive-subject/> (accessed

February 15, 2013).

Zickuhr, K. 2010. *Generations 2010*. Pew Research Center.

Zimmermann, J., J. Forlizzi, and Sh. Evenson. 2006. "Research through Design as a Method for Interaction Design Research in HCI." In *Proceedings of SIGCHI*, 493-502.

Zimmerman, J., and J. Forlizzi. 2008. "The Role of Design Artifacts in Design Theory Construction." In *Human-Computer Interaction Institute*, Paper 37.

Appendix 1: NHS PSA decision aid

Deciding whether to have a prostate specific antigen (PSA) test

This short decision aid is for men who are considering having a test to find out more about their risk of having prostate cancer. The decision aid is not for people who have symptoms of prostate cancer and need a test to diagnose cancer.

The prostate specific antigen (PSA) test can tell you if you have a raised chance of prostate cancer. The PSA test cannot tell you for certain whether you have prostate cancer. If you decide to have a PSA test you can discuss this with your GP, who can arrange for you to have a test.

Whether you decide to have a test will depend on many things, including whether there is a history of prostate cancer in your family, whether you want to know about a raised risk of cancer, and what you think you would do if you had a result that showed a raised chance of prostate cancer.

The options are:

- Have a blood test for prostate specific antigen
- Do not have a blood test for prostate specific antigen.

What are my options?

	Having the PSA test	Not having the PSA test
What is the choice?	<p>The PSA blood test can help to diagnose prostate cancer. It measures the amount of a protein called prostate-specific antigen (or PSA) in your blood. PSA is made by the prostate gland. It is normally found in semen, the fluid which contains your sperm. If there is cancer in the prostate, more PSA leaks into the blood. The higher the level of PSA in the blood sample, the more likely it is to be a sign of cancer.</p> <p>A PSA test on its own cannot tell you for certain whether you have prostate cancer. It can only tell you if you have a raised risk of prostate cancer. PSA levels can vary between men of the same age. Other illnesses which are not cancer can also cause a rise in PSA. If you have a raised PSA level you will need to have other tests to find out if this is caused by prostate cancer.</p>	<p>If you choose not to have a PSA test, you don't have a blood test to find out whether you have a raised risk of prostate cancer.</p>

	Having the PSA test	Not having the PSA test
What is the effect on how long you live?	<p><u>On average</u>, men who have a PSA test do not live longer or shorter lives than men who do not have a test.[1]</p> <p>This information comes from studies looking at <u>big groups of men</u> who were offered PSA testing. <u>We don't know whether having a PSA test will affect your individual length of life.</u></p>	<p>On average, men who do not have a PSA test do not live longer or shorter lives than men who do have a test.[2]</p> <p>This information comes from studies looking at big groups of men who were offered PSA testing. We don't know whether not having a PSA test will affect your individual length of life.</p>

	Having the PSA test	Not having the PSA test
What is the effect on chances of dying from prostate cancer?	<p>Men who have prostate cancer are less likely to die of prostate cancer if they had a PSA test. Their chances of dying of prostate cancer reduce by about 1 in 1000.[3]</p> <p>We don't know if having a PSA test means you're more likely to live longer if you are diagnosed with prostate cancer. [4] This is because having a PSA test does not affect how likely you are to die early of other causes.</p> <p>When all men in a big group are screened, studies show that 293 men need to have a PSA test to prevent one man dying from prostate cancer.[5]</p>	<p>Men who have prostate cancer are more likely to die of prostate cancer if they didn't have a PSA test, compared to men who did have a test. Their chances of dying of prostate cancer are about 1 in 1000 higher.[6]</p> <p>We don't know if not having a PSA test means you're more likely to live longer if you are diagnosed with prostate cancer. [7]</p> <p>Not having a PSA test does not prevent death from prostate cancer.[8]</p>

	Having the PSA test	Not having the PSA test
What is the effect on chances of being diagnosed with prostate cancer?	More men who have a PSA test find out they have prostate cancer than those who don't have a test. Between 6 and 7 in 100 men who have a PSA test are diagnosed with prostate cancer. Between 4 and 5 in 100 men who don't have a PSA test are diagnosed with prostate cancer.[9]	Fewer men who don't have a PSA test find out they have prostate cancer than those who do have a test. Between 4 and 5 in 100 men who don't have a PSA test are diagnosed with prostate cancer. Between 6 and 7 in 100 men who have a PSA test are diagnosed with prostate cancer.[10]

	Having the PSA test	Not having the PSA test
What is the effect on chances of being diagnosed early?	Prostate cancer is diagnosed six to eight years earlier in men who have a PSA test.[11]	Prostate cancer is diagnosed six to eight years later in men who don't have a PSA test.[12]

	Having the PSA test	Not having the PSA test
What is the effect of being diagnosed and treated early?	If you have the PSA test, you're more likely to be diagnosed and treated early. We're not sure if early diagnosis and treatment makes a difference to how long you are likely to live. Not many studies have compared different treatments to no treatment. It also depends on your type of cancer and how advanced it is.	If you don't have the PSA test, you are not likely to be diagnosed and treated early. We're not sure if early diagnosis and treatment makes a difference to how long you are likely to live. Not many studies have compared different treatments to no treatment. It also depends on your type of cancer and how advanced it is.

	Having the PSA test	Not having the PSA test
What are the chances of having a positive test result when you don't have cancer (a false positive)?	The PSA test can suggest you have cancer when you don't (overdiagnose cancer). Most men who have a raised PSA level don't have prostate cancer. Between 65 in 100 [13] and 76 in 100 [14] men who have a raised PSA result do not have prostate cancer, when they have a biopsy test to check. Doctors call this a false positive result.	If you don't have a PSA test, you won't get a false positive result.

	Having the PSA test	Not having the PSA test
What are the chances of having a negative test result when you do have cancer (a false negative)?	The PSA test can miss cases of cancer (underdiagnose cancer). Around 15 in 100 men who have a normal PSA level when they are tested will have prostate cancer.[15] Doctors call this a false negative result.	If you don't have a PSA test, you won't get a false negative result.

	Having the PSA test	Not having the PSA test
<p>What is the effect on chances of needing a biopsy and having complications from biopsy or treatment?</p>	<p>In one large group of men, between 16 and 17 in 100 men who had a PSA test had a result showing a raised PSA level. [16] If the results of your PSA test show you have a raised PSA level, your doctor is likely to suggest you have a biopsy. About 86 in 100 men with a raised PSA level after a PSA test will have a biopsy. [17]</p> <p>If you are diagnosed with prostate cancer, you will need to decide about treatment. Some men have problems or complications after a biopsy or treatment for prostate cancer. Problems from prostate cancer treatment include urinary incontinence and difficulty getting an erection. These problems may not go away.</p>	<p>If you don't have a PSA test, it's unlikely you will need to have a biopsy unless you get symptoms of prostate cancer. This means you are unlikely to have any complications from biopsies or treatments.</p>

What are the pros and cons of each option?

People have different views on the pros and cons of having a test for prostate specific antigen (PSA) to find out more about their risk of having prostate cancer. Choosing the option that is best for you means considering how the consequences of each option - having a blood test or not having a blood test - will affect your life.

Here are some questions people may want to consider before deciding whether have the PSA test.

- Do they only want to take the test if it will lower their chances of dying from prostate cancer?
- Do they want to know for sure whether or not they have prostate cancer?
- Are they willing to take the PSA test if it's likely they will need more tests afterwards?
- Are they willing to take the PSA test if it's likely they will need treatments afterwards?
- How important is it to them that the test is reliable?
- Are they willing to take the PSA test if there's a chance they will have side effects or problems?

How do I get support to help me make a decision that is right for me?

Go to <http://sdm.rightcare.nhs.uk/pda/psa-testing/> for more detailed information about treatments for **Prostate Specific Antigen (PSA) Testing**. People using this type of information say they understand the health problem and treatment choices more clearly, and why one treatment is better for them than another. They also say they can talk more confidently about their reasons for liking or not liking an option with health professionals, friends and family.

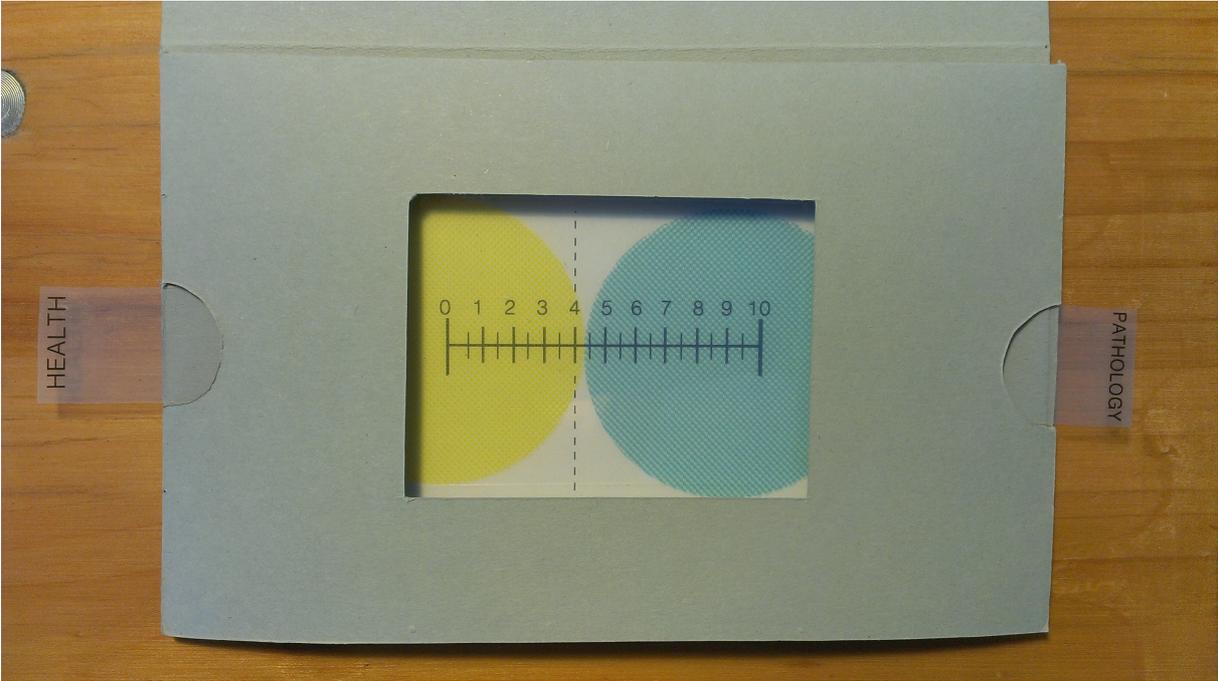
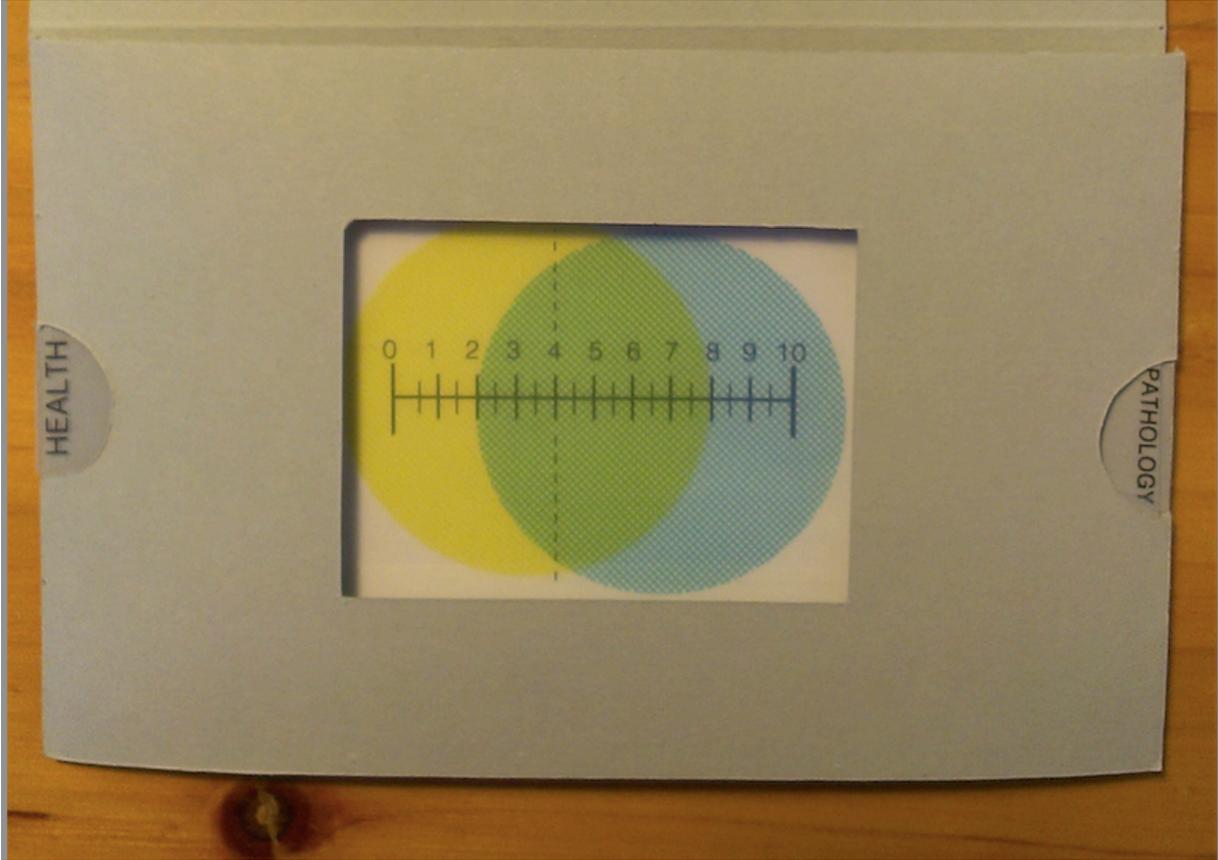
You can call our Decision Support service on 0845 450 5851 to speak to a trained Health Coach. The Health Coaches will assist you by:

- Providing you with information
- Helping you to understand your condition
- Recognise what is important to you regarding the outcome of treatment
- Identifying potential solutions
- Encourage you in discussing options with your family
- Transferring skills which will assist you in using the information and resources available to you
- Support you in building confidence in discussing your choice with your doctor
- You may find that this can be achieved in one telephone call with a Health Coach; however, if further calls are required to support you in reaching your decision the Health Coach will schedule these with you.

References

References can be viewed online at <http://sdm.rightcare.nhs.uk/pda/psa-testing/references/>

Appendix 2: Prototype 1 – Venn Interaction Diagram



Appendix 3: Prototype 2 – Reflection cards

du und dein Arzt sprechen über die nächsten Schritte

Du hattest einen PSA-Bluttest mit einem Ergebnis leicht über dem Grenzwert. Dein Arzt empfiehlt dir eine Gewebeentnahme (Biopsie) vorzunehmen, um sicher zu gehen. Du spielst mit dem Gedanken.

Wie reagierst du, wenn dir am Abend eine Freundin entgegnet, dass PSA-Bluttests sehr oft falsch-positiv Ergebnisse angeben?

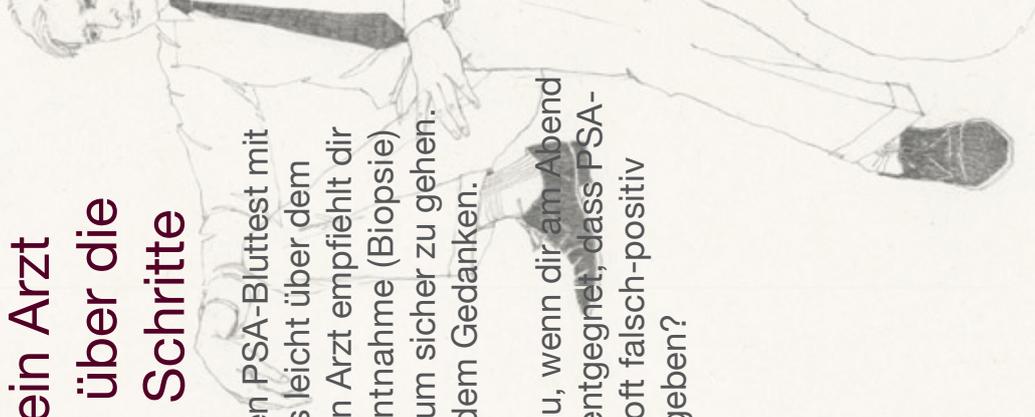
s1_01

in der Liebe

Du hattest einen leicht positiven PSA-Bluttest und dein Arzt schlägt eine Gewebeentnahme (Biopsie) vor. Einen solchen Eingriff vorzunehmen beunruhigt dich. Bei deinem Freund, der regelmässig Biopsien vornehmen lässt, wurde nie etwas gefunden. Dein Partner fragt dich, ob du nicht auch Gewissheit haben möchtest.

Wie reagierst du?

s1_02



Im ersten Moment denke ich...

.....

.....

.....

.....

.....

.....

.....

Im ersten Moment denke ich...

.....

.....

.....

.....

.....

.....

.....

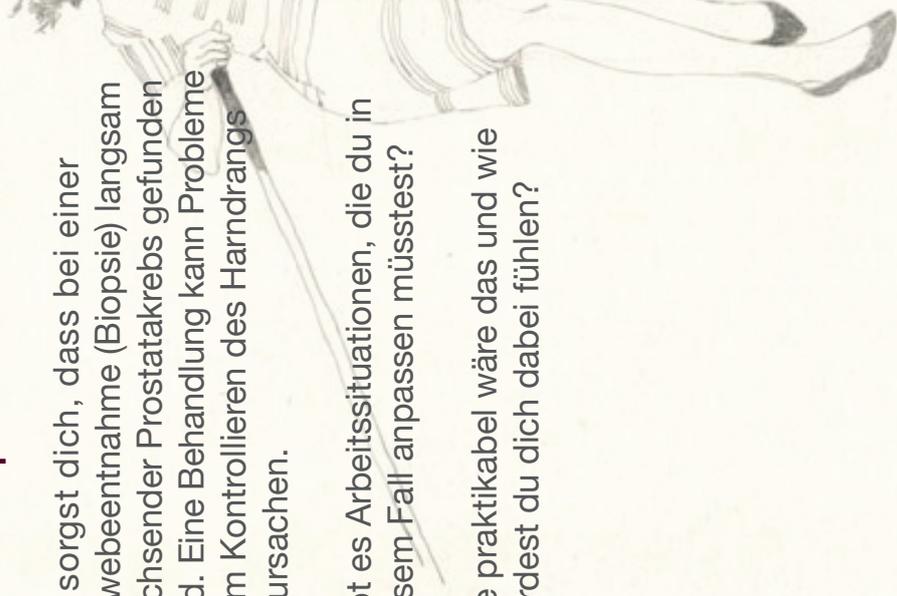
wie flexibel ist dein Arbeitsplatz

Du sorgst dich, dass bei einer Gewebeentnahme (Biopsie) langsam wachsender Prostatakrebs gefunden wird. Eine Behandlung kann Probleme beim Kontrollieren des Harndrangs verursachen.

Gibt es Arbeitssituationen, die du in diesem Fall anpassen müsstest?

Wie praktikabel wäre das und wie würdest du dich dabei fühlen?

s1_03



in der Liebe

Du hattest ein positives Ergebnis beim PSA-Bluttest. In einem Magazin liest du, dass eine präventive Prostatatektomiebehandlung möglicherweise Erektionsprobleme zur Folge haben kann. Du überlegst, welche Rolle Sex in deinem Leben spielt.

Wie würde sich dein Partner fühlen, wenn du Erektionsprobleme hättest?

s1_04



Ich denke ...

.....

.....

.....

.....

.....

.....

.....



Ich denke ...

.....

.....

.....

.....

.....

.....

.....

unter Kollegen

Du hattest einen positiven PSA-Bluttest und hast dich für eine Gewebeentnahme (Biopsie) entschieden. Während des Mittagessens auf der Arbeit hörst du jemanden über einen falschen Biopsie-Befund sprechen.

Wie fühlst du dich dabei?

s2_05



Hobbies mit deinem Partner

In zehn Tagen lässt du dir eine Gewebeentnahme entnehmen. Du bist besorgt, dass man dir möglicherweise eine Prostatakrebsbehandlung empfiehlt. Dies kann Inkontinenz (Probleme den Harndrang zu kontrollieren) zur Folge haben. Dein Partner bucht gerade Flüge für eure nächsten Ferien.

Wie würde Inkontinenz eure Hobbies und Ferien beeinflussen?

s2_06

Was denkt dein Partner?

.....

.....

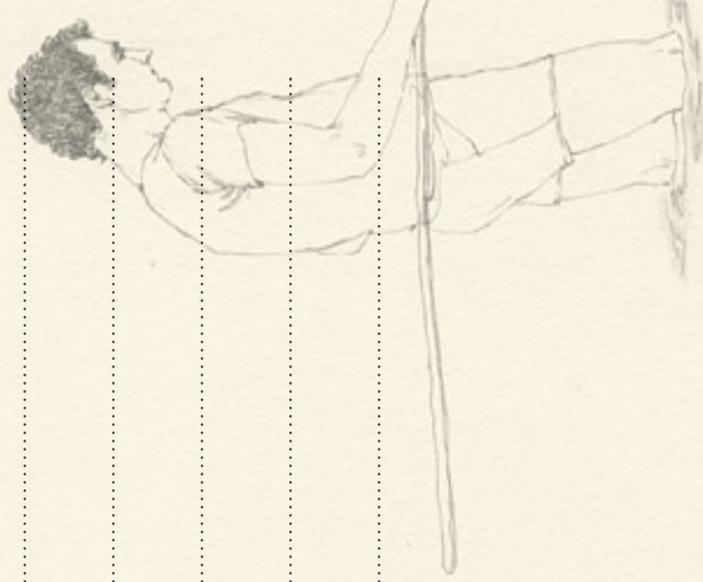
.....

.....

.....

.....

.....



Was unternimmst du?

.....

.....

.....

.....

.....

.....

.....

du und dein Arzt besprechen nächste Schritte

Du hast besorgt auf das Resultat deiner ersten Gewebeentnahme (Biopsie) gewartet und soeben ein negatives Resultat erhalten. Trotzdem empfiehlt dein vertrauter Arzt aktive Überwachung, PSA-Bluttests alle sechs Monate und weitere Biopsien wenn nötig.

Wie reagierst du und wieso?

s3_07

unter Freunden

Du bist erleichtert. Das Resultat deiner ersten Gewebeentnahme (Biopsie) ist negativ. Beim Feierabendbier erzählt dir ein Freund, dass diese Befunde auch oft fehlerhaft sein können.

Wie reagierst du und was unternimmst du?



s3_08

Ich denke ...

.....

.....

.....

.....

.....

.....

.....

Fühlst du dich wohl mit so einem
Vorgehen?

.....

.....

.....

.....

.....

.....

.....



unter Freunden

Der Befund deiner ersten Gewebeentnahme (Biopsie) ist negativ. Dein vertrauter Arzt empfiehlt dir mit den Tests fortzufahren. Eine gute Freundin rät dir davon ab. Sie erzählt dir, dass ihr Vater trotz häufiger Biopsien nie bösartigen Prostatakrebs hatte.

Wie reagierst du?

s3_09

in der Familie

Du bist erleichtert. Du hattest eben deine Gewebeentnahme (Biopsie) und erneut ein negatives Resultat. Dein Schwager hat langsam wachsenden Prostatakrebs und rät dir mit dem Test fortzufahren.

Beeinflusst dich das?

s3_10

Wieso, oder wieso nicht?

.....

.....

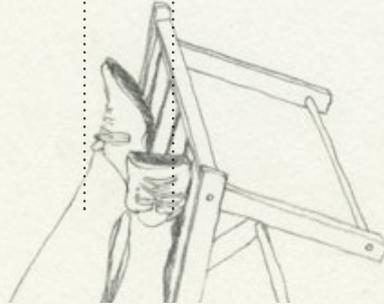
.....

.....

.....

.....

.....



Mit wem sprichst du darüber?

.....

.....

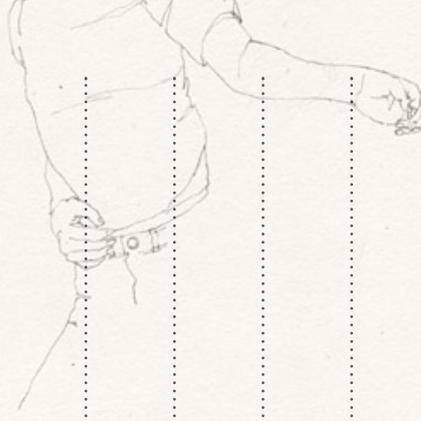
.....

.....

.....

.....

.....



im Gespräch mit deinem Partner

Du bist erleichtert. Du hattest eben eine Gewebeentnahme (Biopsie) und erneut einen negativen Befund. Deine Partnerin ist skeptisch, ob das Resultat stimmt und rät dir mit den Tests fortzufahren.

Wie fühlst du dich dabei?



s3_11

in der Liebe

Deine erste Gewebeentnahme (Biopsie) vor zwei Jahren war negativ. Dein Arzt empfiehlt dir mit den Tests fortzufahren, aber du lehnt ab. Deine neue Partnerin fragt dich, wie du mit diesem Thema umgehst.

Wie erklärst du deine Einstellung?

s3_12

Meiner Meinung nach ...

.....

.....

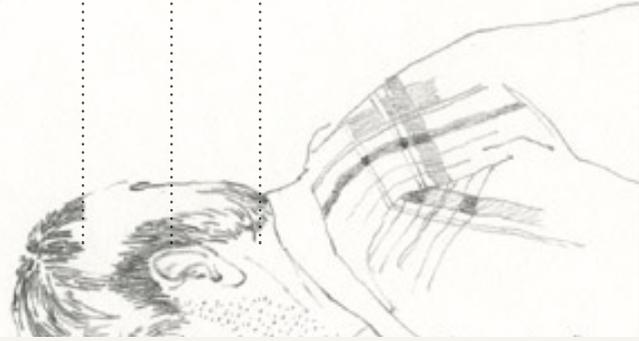
.....

.....

.....

.....

.....



Was unternimmst du?

.....

.....

.....

.....

.....

.....

.....

unter Freunden

Du hattest eben deine erste Gewebeentnahme (Biopsie) und ein negatives Resultat. Allerdings hast du nun leichte Schmerzen beim Urinieren. Dein Arzt empfiehlt dir weiterhin Blut- und Gewebetests vorzunehmen.

Was wirst du deinem Freund empfehlen, der sich entscheiden muss, ob er dem gleichen Verfahren zustimmt oder nicht?

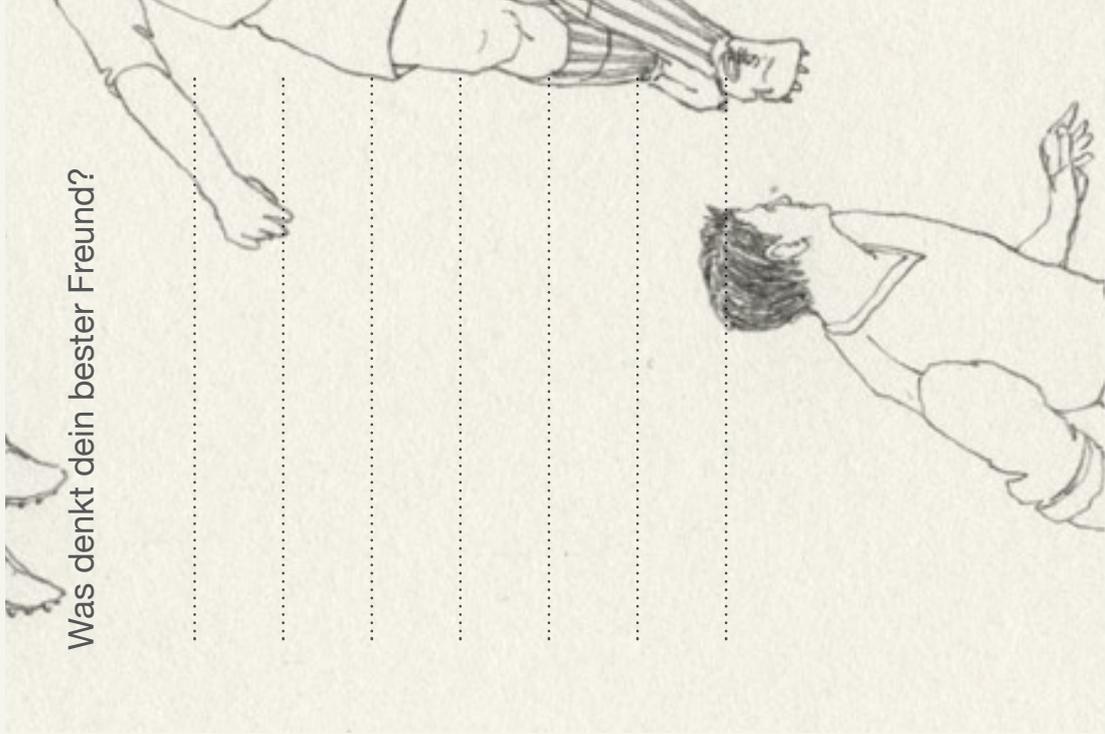
s3_13

unter Freunden

Du hattest eben deine zweite Gewebeentnahme (Biopsie) und ein negatives Resultat, aber du leidest an temporären Nebeneffekten. Du machst dir Gedanken über die möglichen Nebeneffekte einer präventiven Prostatatektomiebehandlung.

Wie würde sich Inkontinenz (Probleme beim Kontrollieren des Harnstrahls) auf deinen Sport auswirken?

s3_14



Ich denke ...

Seven horizontal dotted lines for writing.

bei der Arbeit

Deine dritte Gewebeentnahme (Biopsie) zeigt erneut ein negatives Resultat. Der Eingriff hat temporär schmerzhaft Nebeneffekte zur Folge. Du machst dir Gedanken über die möglichen Nebeneffekte einer präventiven Prostatakrebsbehandlung.

Wie würde sich Inkontinenz (Probleme beim Kontrollieren des Harnstrahls) auf deinen Beruf auswirken?

Was denkt dein Kollege?

s3_15

mit deiner Familie sprechen

Deine dritte Gewebeentnahme (Biopsie) zeigt ein negatives Resultat. Der Eingriff hat temporär eine schmerzhaft Entzündung zur Folge. Du machst dir Gedanken über eine präventive Prostatakrebsbehandlung.

Könntest du offen mit deiner Familie darüber sprechen?

s3_16

Was denkt dein Partner?

.....

.....

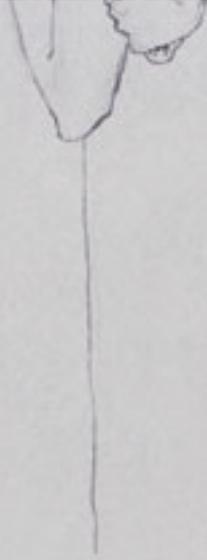
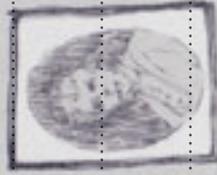
.....

.....

.....

.....

.....



Was meint dein Vorgesetzter?

.....

.....

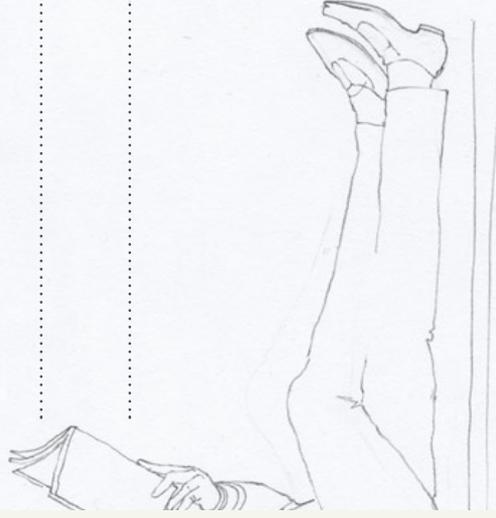
.....

.....

.....

.....

.....



dein Arbeitsplatz

Nach deiner dritten Gewebeentnahme (Biopsie) ist das Resultat wiederholt negativ. Aber du hast kurzzeitig Schmerzen beim Urinieren. Während eines Arbeitsmeetings musst du darüber nachdenken, wie du mit Inkontinenz (Kontrollverlust des Harndrangs) am Arbeitsplatz umgehen würdest. Inkontinenz kann als Nebeneffekt einer Prostatakrebsbehandlung auftreten.

Ist es möglich kurz vor deiner Beförderung offen damit umzugehen?

s3_17



du und dein Partner sprechen über nächste Schritte

Du erwachst aus der Narkose nach einer Gewebeentnahme (Biopsie). Deine Urologin teilt dir mit, dass du mittel-aggressiven Prostatakrebs hast. Sie empfiehlt dir eine komplette Entfernung der Prostata. Dies kann Erektionsprobleme zur Folge haben. Zuhause wartet dein Partner.

Wie fühlst du dich?

s4_18

Habt ihr bereits offen über sexuelle
Wünsche gesprochen?

.....

.....

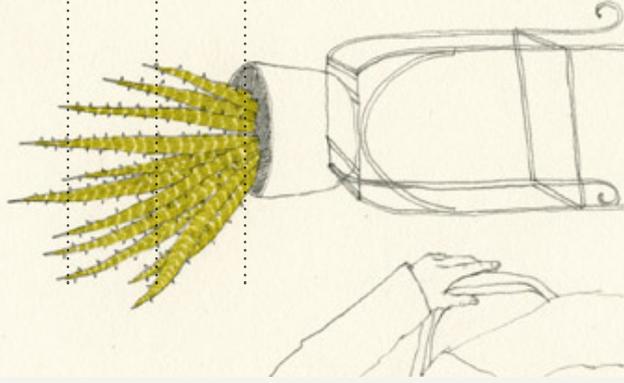
.....

.....

.....

.....

.....



Ich denke ...

.....

.....

.....

.....

.....

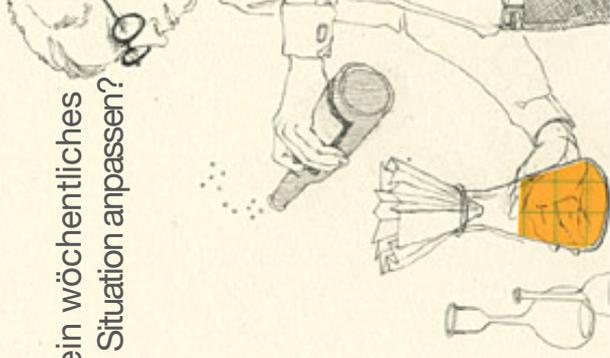
.....

.....

unter Sportskameraden

Nach einer Gewebeentnahme (Biopsie) hast du temporär Probleme den Harnrang zu kontrollieren. Dein Urologe teilt dir mit, dass du aggressiven Prostatakrebs hast und empfiehlt eine zügige Behandlung. Du bist besorgt, dass sich das Harnrangproblem verschlimmern wird.

Wie könntest du dein wöchentliches Training einer solchen Situation anpassen?



s5_19

beim Essen mit deiner Familie

Du hattest eben eine Gewebeentnahme (Biopsie) mit Diagnose langsam wachsender Prostatakrebs. Dein Arzt empfiehlt aktive Überwachung mit Bluttests alle sechs Monate.

Was wirst du deiner Familie beim Abendessen berichten und wie wirst du dich entscheiden?

s6_20

Wieso?

.....

.....

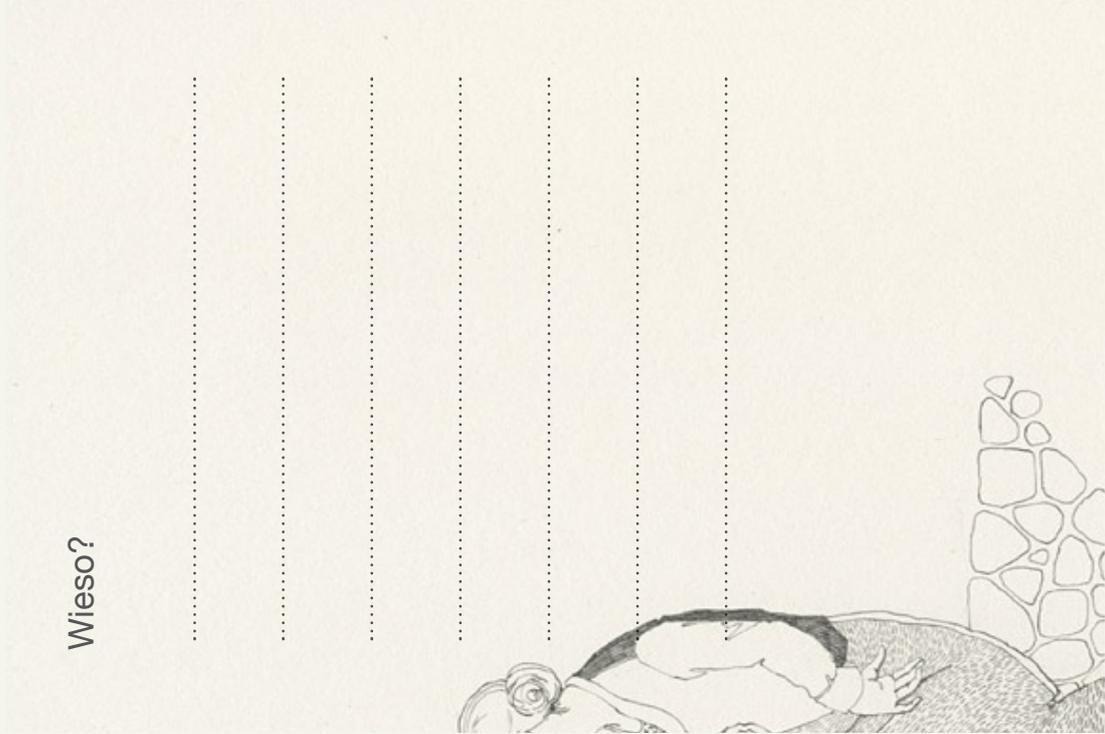
.....

.....

.....

.....

.....



Ich finde ...

.....

.....

.....

.....

.....

.....

.....

unter Kollegen

Nach einer Gewebeentnahme (Biopsie) mit Diagnose langsam wachsender Prostatakrebs entscheidest du dich für watchful waiting. Bei deinem Vorgesetzten wurde gerade die Prostata entfernt, weil man aggressiven Krebs gefunden hat.

Wie erklärst du deinem Kollegen deine Entscheidung?



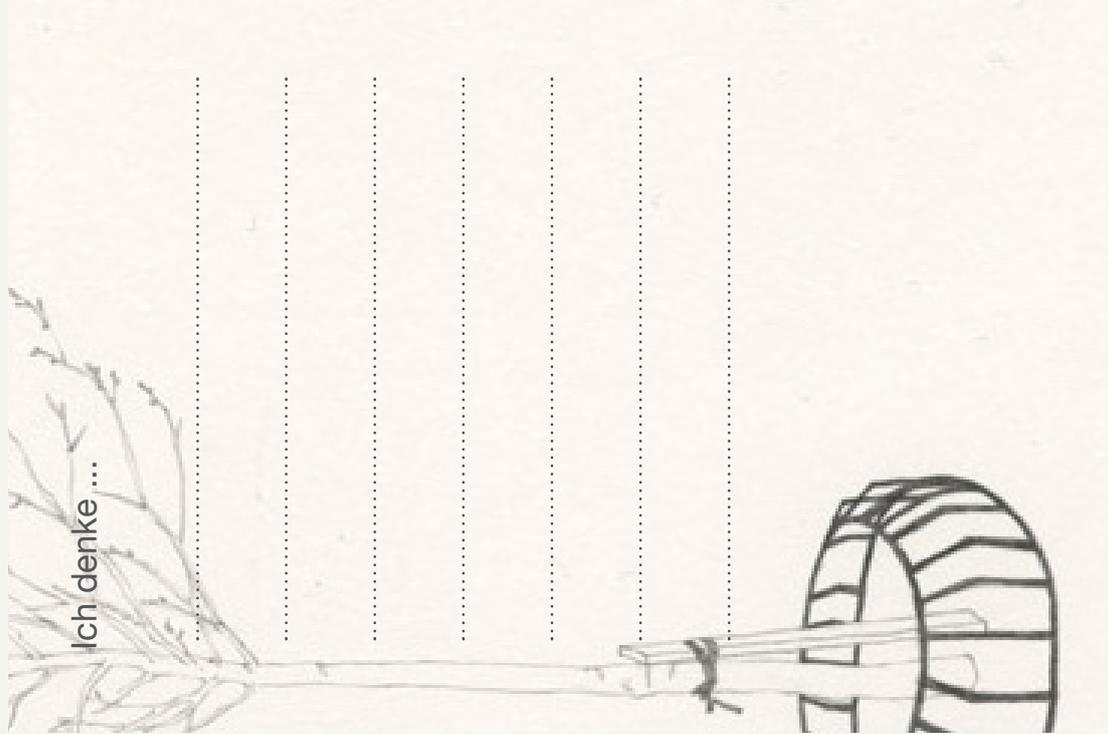
s6_21

unter Freunden

Dir wurde eben eine Gewebeprobe entnommen. Der Befund deutet auf langsam wachsenden Krebs hin. Du entscheidest dich für aktive Überwachung. Dein nächster PSA-Bluttest ist in neun Monaten.

Wie beschreibst du deine Situation wenn deine Freunde wissen wollen, wie es dir gesundheitlich geht?

s6_22



Ich finde ...

Seven vertical dotted lines for writing.

du und deine Familie

Du hattest einen positiven PSA-Bluttest und deine Urologin empfiehlt eine Gewebeprobe zu entnehmen. Nach langem Überlegen entscheidest du dich dagegen. Dein Onkel erzählt dir, dass seine Nachbarin wegen zu spätem Feststellen von Brustkrebs beinahe nicht mehr behandelt werden konnte.

Wie erklärst du deine Entscheidung?

s7_23

unter Freunden

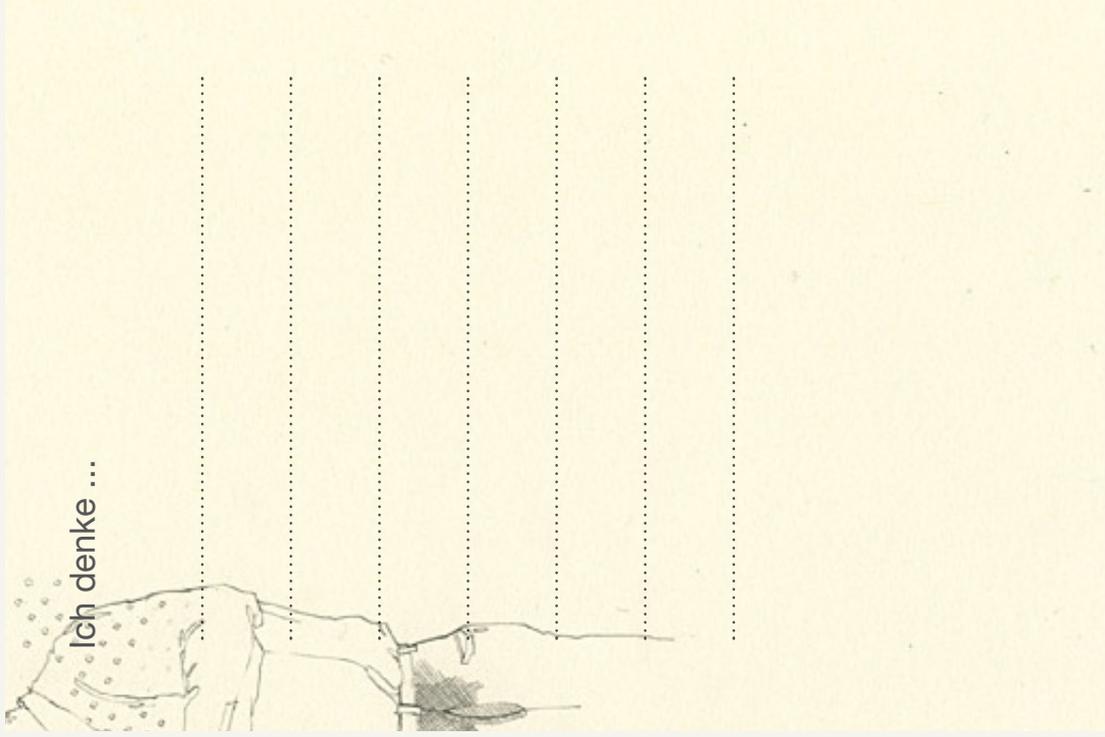
Du hattest einen positiven PSA-Bluttest. Dein Arzt empfiehlt dir eine Gewebeentnahme (Biopsie) aber du lehntst ab.

Wie erklärst du deiner besorgten Familie deine Entscheidung?



s7_24

Ich denke ...



.....

.....

.....

.....

.....

.....

.....

Ich finde ...

.....

.....

.....

.....

.....

.....

.....

in der Liebe

Du hattest eben deine zweite Gewebeentnahme (Biopsie) in drei Jahren und erneut ein negatives Resultat. Dein Arzt empfiehlt mit den Tests fortzufahren. Wiederholtes Testen und mögliche Nebeneffekte beunruhigen dich.

Wie reagierst du wenn deine Partnerin über ihren Kinderwunsch spricht?



Ich denke ...

.....

.....

.....

.....

.....

.....

.....

Prototype 2 - Translation of Reflection Cards

S1_01

You and your GP talk about next steps

You just had a PSA blood test and the result was slightly above the recommended threshold. Your GP recommends conducting a tissue test (biopsy), just to be sure. You are considering doing it.

How do you respond when you meet a friend in the evening and she says that PSA blood tests very often show false-positive test results?

S1_02

Concerning love

You had a slightly positive PSA blood test and your GP recommends conducting a tissue test (biopsy). You are worried about conducting such an intervention. A friend has already conducted a few biopsies but no cancer was ever found.

Your partner asks you whether you would not rather have certainty about this. How will you respond?

S1_03

How adaptable is your work place?

You are worried that conducting a tissue test (biopsy) would find slow-growing prostate cancer. Preventive treatment can cause problems with urination.

Are there work-related situations which you would need to modify as a result of having problems with controlling urination? How feasible would that be and how would you feel about it?

S1_04

Concerning love

You had a positive PSA blood test. In a magazine, you read that preventive prostate cancer treatment may cause erection problems.

You think about the role sex plays in your life. How would your partner feel like in case you had erection problems?

S2_05

Among work colleagues

You had a positive PSA blood test and decided to conduct a tissue test (biopsy). During lunch you hear someone talking about a false tissue test result.

How do you feel about that? What will you do?

S2_06

Leisure activities with your partner

In ten days you will conduct a tissue test. You are concerned that preventive prostate cancer treatment will be recommended, and this can cause problems controlling urination.

Your partner is booking flights for the upcoming holidays. In what ways would a problem with controlling urination affect your leisure activities and holidays? What does your partner think about it?

S3_07

You and your GP talk about next steps

You have anxiously awaited the result of your tissue test (biopsy). The result was negative, nevertheless your GP recommends following a watchful waiting approach with PSA blood tests every six months and further tissue tests if necessary.

What is your reaction to this? Do you feel comfortable with such an approach?

S3_08

Among friends

You are relieved. The result of your tissue test (biopsy) is negative. While drinking a beer with a friend in the evening, he mentions that such results can be incorrect.

How do you respond and what will you do?

S3_09

Among friends

The result of your first tissue test (biopsy) is negative. Your trusted GP recommends continuing testing. A good friend discourages you from continuing testing. She mentions that, despite conducting periodic tissue testing, her father never had aggressive prostate cancer.

How do you respond? Who will you talk to about this?

S3_10

Family

You are relieved. You just had a tissue test (biopsy) and again a negative result. Your brother in law has slow-growing prostate cancer and recommends continuing testing.

Will this affect your decision? Why, or why not?

S3_11

In conversation with your partner

You are relieved. You just had a tissue test (biopsy) and again a negative result. Your partner is skeptical about the test result and recommends continuing testing.

How do you feel about it? What will you do?

S3_12

Concerning love

Your first tissue test (biopsy) two years ago was negative. Your GP recommends continuing testing, but you decline.

Your new partner is asking you how you deal with this situation. How do you explain your decision on this?

S3_13

Among friends

You just had your first tissue test (biopsy) with a negative result. However, you now suffer from slight pain when urinating. Your GP recommends continuing blood and tissue testing.

What will you recommend your friend who is about to decide whether or not to screen for prostate cancer?

S3_14

Among friends

You just had your second negative tissue test (biopsy) but suffer from temporary side-effects. You are thinking about the potential side-effects of a preventive prostate cancer treatment.

In what ways would problems with controlling urination affect your sport activities? What does your best friend think about it?

S3_15

At work

Your third tissue test (biopsy) is again negative. However, the intervention led to temporary pain. You are thinking about the potential side-effects of a preventive prostate cancer treatment.

In what ways would problems with controlling urination affect your work? What does your colleague think about it? What about your boss?

S3_16

Family

Your third tissue test (biopsy) is again negative. However, the intervention led to temporary pain. You are thinking about the potential side-effects of a preventive prostate cancer treatment.

Can you talk about it openly with your family? What does your partner think about it?

S3_17

At work

Your third tissue test (biopsy) is again negative. However, the intervention led to temporary pain. During a meeting at work, you are thinking about how problems with controlling urination (incontinence) would affect your work situation. Incontinence can be a potential side-effect of preventive prostate cancer treatment.

Is it possible to handle your concern openly shortly before your expected promotion?

S4_18

Talking about next steps with your partner

You are waking up from anesthesia from having a tissue test (biopsy). Your urologist tells you that you have medium-aggressive prostate cancer. She recommends a complete removal of the prostate gland. This can cause problems having an erection.

At home your partner is waiting for you. How do you feel about it? Have you already discussed sexual desires openly?

S5_19

Among football friends

After a tissue test (biopsy) you suffer from temporary problems controlling urination. Your urologist tells you that you have aggressive prostate cancer and recommends swift treatment. You are also concerned that the problem with controlling urination will get worse.

How would you adapt your weekly football workout to the situation?

S6_20

At the family dinner

You just had a tissue test (biopsy) and been diagnosed with slow-growing prostate cancer. Your GP recommends watchful waiting with blood tests every six months.

What will you tell your family over dinner tonight and how will you decide? Why?

S6_21

Among work colleagues

After a tissue test (biopsy) and a slow-growing prostate cancer diagnosis, you decide to continue with watchful waiting. Your boss just had his prostate removed due to an aggressive prostate cancer diagnosis.

How will you account for your decision with your colleagues at work?

S6_22

Among friends

After a tissue test (biopsy) and a slow-growing prostate cancer diagnosis, you decide to continue with watchful waiting. Your next PSA blood test is in 9 months.

How will you describe the current situation to your friends when they ask you about your health and wellbeing?

S7_23

You and your family

You had a positive PSA blood test and your urologist recommends conducting a tissue test (biopsy). After intense consideration you decide against it. Your uncle mentions that his neighbour could almost not be treated anymore as her breast cancer was found late.

How do you account for your decision?

S7_24

Among friends

You had a positive PSA blood test and your GP recommends conducting a tissue test (biopsy). You decline.

Your family is concerned. How will you explain your decision?

S7_25

Concerning love

You just had your second negative tissue test (biopsy) in three years. Your GP recommends continuing testing. Periodic testing and the potential side-effects worry you.

How will you respond to your partner when she raises her wish to start a family?

