Summary Report of Qualitative Research into Public Attitudes to Personal Data and Linking Personal Data

July 2013

wellcometrust

THE WELLCOME TRUST

SUMMARY REPORT OF QUALITATIVE RESEARCH INTO PUBLIC ATTITUDES

TO PERSONAL DATA AND LINKING PERSONAL DATA

July 2013

A. RESEARCH BACKGROUND & OBJECTIVES

The Wellcome Trust commissioned CM Insight to conduct qualitative research to understand the general public's attitudes to different types of personal data and data linking. The research looked at whether health data are viewed differently from other types of data, and what are the perceived risks and benefits, to self and society, of linking different kinds of data for research purposes and other purposes.

B. METHOD & SAMPLE

Six 90-minute focus groups and six 45-minute telephone interviews were conducted with members of the public, between 29 April and 12 May 2013. A total of 50 people took part. The sample included men and woman, aged 18-70 from socio-economic groups ABC1 and C2DE. The focus group respondents were recruited as owners of products such as a loyalty store card, credit card, smart phone and driving licence; some shopped or banked online, and many used social media. Telephone interviewees were defined as the 'pro-privacy group': they were recruited as especially cautious about releasing personal data and actively taking some measures to protect against doing so, such as ensuring they opted out of having their details passed onto third parties, shredding correspondence with their name and address on it, and never shopping online or by phone using a credit card. All respondents completed a pre-task (self-completion questionnaire) before their research group/interview, which involved identifying benefits and disadvantages about personal data collection including health data, and asking these questions of two people they knew. The research took place in London, Midlands and Norfolk, representing urban, suburban and rural areas respectively.

C. MAIN FINDINGS

1. Context

It is helpful to consider responses to data and data linkage in the context of life today. The general public has an awareness of powerful external forces, namely: a Government that is seeking to reduce spending in the economic crisis; a myriad of fraudsters and cheats, swindling the State and the individual; aggressive and increasingly sophisticated individual marketing; and a fast-changing technological world that is hard for many people to keep up with, plus a general push towards everything moving online. Trust has been lost in major institutions (e.g. banks) and the NHS is in crisis, with some fearing discrimination in provision of health services in future. People feel watched by Big Brother as well as fraudsters; there are stories of data security breeches in the press, as well as snooping by the press. It is not surprising, therefore, that cynicism towards 'them' (the Government, corporations and press) and a significant fear of being a victim of fraud were observed.

2. Perceptions of personal data

2.1 Sample differences

In terms of concern about personal data collection/use, there was little obvious difference between men and women. There were observable differences, however, by age and social grade, with young people seeming more accepting of data collection/use than older people, and C2DEs feeling more powerless to deal with consequences, e.g. arguing their case if their identity were stolen.

The pro-privacy respondents were generally more fearful of 'what ifs', with the majority having a higher awareness of issues and difficulties that could arise. However there was also some overlap in attitude/behaviour between the focus group participants and pro-privacy sub-group, suggestive of a spectrum of views generally, rather than people falling into two discrete groups. Some pro-privacy respondents had been victims of fraud, or knew people who had been. Some were cautious in releasing personal data because they worked in industries responsible for handling personal data (e.g. banking, insurance, selling databases). One middle-aged

respondent (not using the internet) was unnerved by a fast-changing technological world she could not keep up with.

2.2 Benefits of personal data collection/use

The main benefits identified were: the Government identifying needs, planning resources and services, and allocating funds; prevention and detection of crime, including terrorism; tailored marketing; identifying social/population trends and statistics; convenience and time-saving when shopping and doing other transactions, if personal data were already held; unearthing dishonesty (e.g. fraudulent benefit claimants and tradesmen); and availability of vital medical information in a medical emergency.

2.3 Drawbacks of personal data collection/use

The main drawbacks identified were: the potential for data to be lost, stolen, hacked and leaked, and shared without consent, leading to security concerns; invasion of privacy, with a sense of Big Brother watching; unsolicited marketing and advertising, with a nigh nuisance factor; incorrect/inaccurate data collection, which would be hard to correct and undo; potential discrimination (e.g. data falling into the hands of a prospective employer).
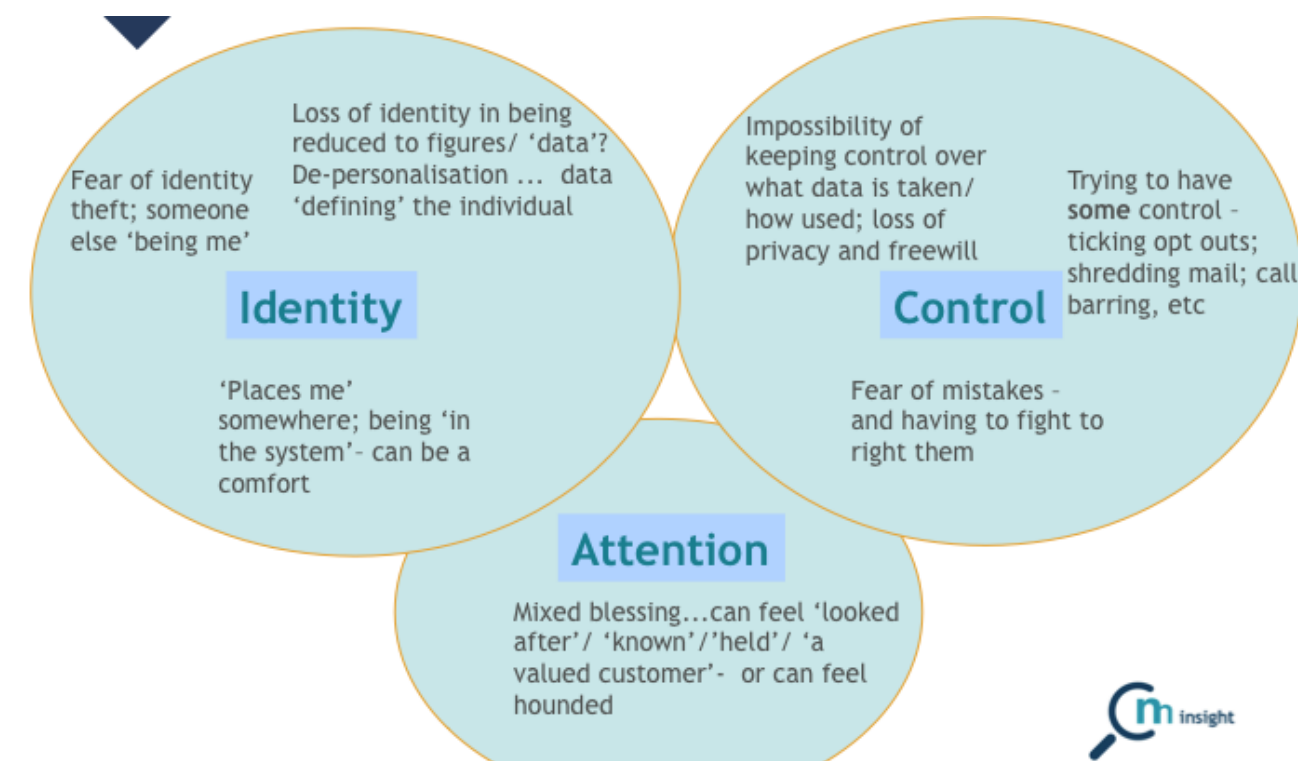
2.4 Different views about health data

Many, though not all, did regard health data differently from other types of data. Some stressed the importance of health data collection when it benefits the individual; indeed, sharing data **within** the NHS was a definite positive (e.g. sharing an individual's health data across different hospitals) and the perception was that more data sharing could usefully be done within the NHS. Yet there was also a strong feeling that personal health data are confidential, private and sensitive, and should not be shared outside secure, authorised bodies such as the NHS, and especially not with private companies such as employers, insurance providers and drug manufacturers. Mental health data was sometimes regarded as particularly personal and sensitive.

Additionally, ABC1s were more likely than C2DEs to view the use of health data as having a potential benefit to society, in the fields of research, disease prevention, planning of services, crime prevention and so on.

2.5 Major issues underlying views on personal data

Three major themes emerged as underlying the public's views on personal data. These are represented in the chart below.

Fear of identity theft; someone else 'being me'

Loss of identity in being reduced to figures/ 'data'? De-personalisation ... data 'defining' the individual

Impossibility of keeping control over what data is taken/ how used; loss of privacy and freewill

Trying to have **some** control – ticking opt outs; shredding mail; call barring, etc

Identity

Control

'Places me' somewhere; being 'in the system'– can be a comfort

Fear of mistakes – and having to fight to right them

Attention

Mixed blessing...can feel 'looked after'/ 'known'/'held'/ 'a valued customer'– or can feel hounded

m insight

2.6 Personal data examples

Focus group respondents were given a set of cards containing over 20 different types of data, and asked to cluster the data types into smaller groups using any criteria they liked. Awareness of the data types appeared latent, i.e. most were aware of the different types of data but did not consciously regard them all as 'personal

data'. This exercise yielded a variety of clusterings of the data types, demonstrating that 'categories' of data, as perceived by the public, are fluid/overlapping.

A qualitative amalgamation/analysis of the clusterings revealed the following key underlying criteria used to distinguish between the data: the degree of seriousness/risk if the data were misused or stolen; the perceived level of security of the data; anonymous vs. personally identifiable data; recognition of the value of data collection (to self vs. to others) vs. unclear benefit; free choice to create data vs. enforced/necessary existence of the data; government and non-government data.

The following chart is an attempt to plot the data on two key dimensions: how compulsory the giving of the data is, in today's world, and level of potential concern about loss/misuse of the data (incorporating both likelihood of it happening, and how serious the consequences if it did).
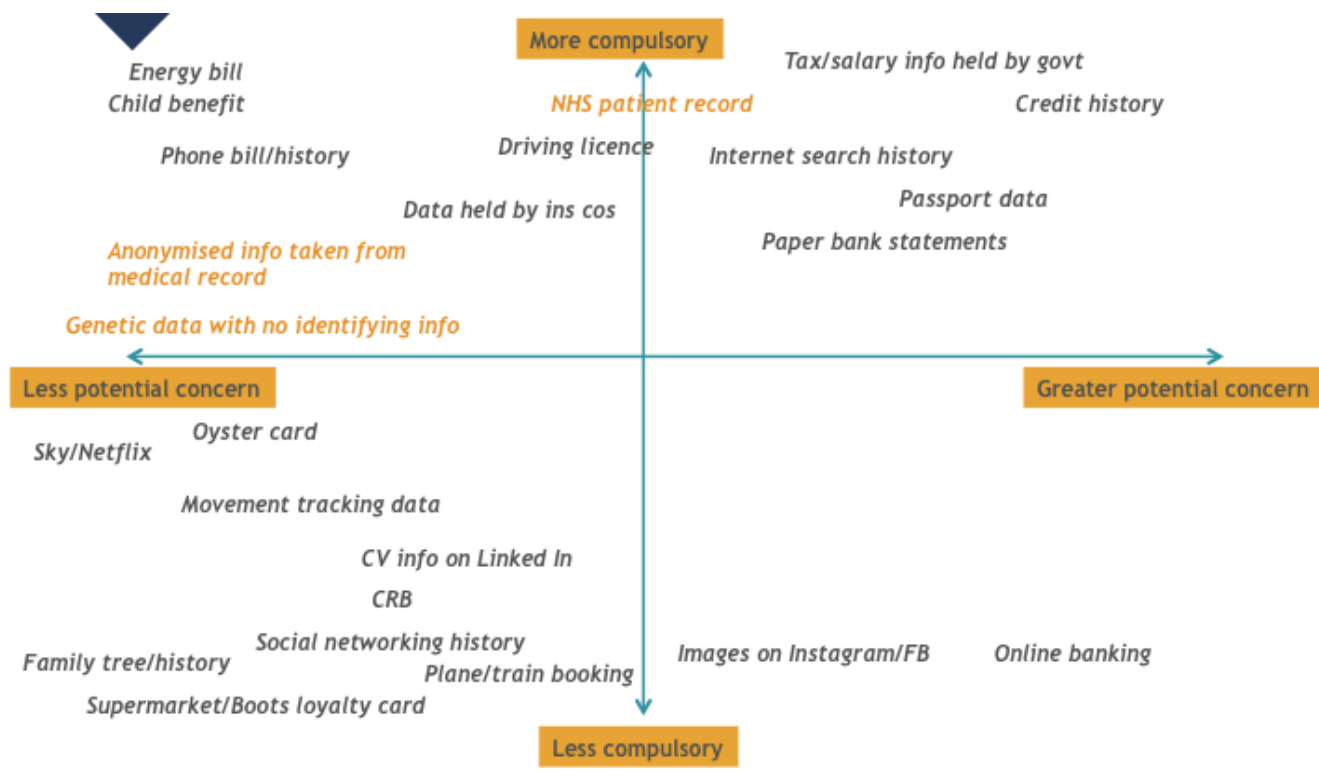
3. Health data

3.1 Different types of health data

Personal health data (e.g. NHS patient health record, test results, dental and optical records) were top of mind, rather than population health data. With personal health data representing protection, care and cure at individual level, especially in circumstances when health is failing, it was reassuring, comforting and necessary for personal health data to be collected by and held within the NHS. A cause for concern, however, was the notion that future discrimination might be introduced, or was already being practiced, within the NHS (e.g. low priority on waiting list; being refused treatment until lifestyle changes are made, such as stopping smoking). Some expressed fear that if parts of the NHS are privatised, then health data may be vulnerable to outside parties. A couple of anecdotal data breeches within the NHS were mentioned (e.g. a list of individuals, including addresses, with STIs found in one respondent's maternity folder) although faith in the NHS had not been greatly shaken as a result.

The thought of personal health data getting into the 'wrong hands' outside the NHS did worry some people, especially if they had a particular history/condition. Examples of the wrong hands and inappropriate use were commercial companies for advertising purposes (e.g. targeting older/frail people with health products), and sharing personal data with employers, who might discriminate against individuals, e.g. because of mental health problems, or HIV status.

Population health data were regarded as anonymous, of benefit to all, and reassuring to be collected for the common good, especially in the long term. However, the possibility of individual identification was a cause for concern, if – again - those data should get into the 'wrong hands' (particularly the media) and be used against an individual.

Health data used in clinical trials and research were much less salient. Whilst there was awareness of research being conducted to find cures and treatments generally, with implicit benefit to individuals, families and future generations, there was low

awareness of ongoing collection of data to establish norms. Awareness of research tended to be vague, with the exception of cancer research generally, and a handful of specific studies (e.g. two Midlands respondents were Biobank volunteers). Fears around research/clinical trial data were low level, and, on prompting, related to anonymity being lost and possible unwanted media attention.

3.2 Examples of health data
Seven different types of health data (see below) were explored in the research; overall, few objections were raised to any. The examples largely served to reassure because they seemed to be in the public interest or for the good of society; future good would come from obtaining this information and understanding. Provided health data are anonymous, used and kept safe within 'the system', i.e. their respective research/health environment, then people have no or little issue. Where there was concern, this related to confidentiality and security of data with fears centring round the media accessing confidential information.

The most clearly beneficial examples of health data collection/use were: commissioning (GP surveillance of chickenpox), public health (GPs reporting measles to health board), and audit (using medical records to measure/compare hospital mortality rates). The sexual health (national survey - NATSAL) and mental health (UK biobank – long term study) examples were acceptable as participation in each was voluntary.

The health data 'research' example (investigating incidence of leukaemia in children in relation to proximity of their home to power lines) raised questions, mainly about who the researchers were, who the research is funded by, and whether parental consent is given. This was an example where respondents had some concern about potential identification of individual children by the media, if information was somehow accessible or leaked.

The health data 'genetic' example (research into origins of rare diseases) also had controversial aspects. Whilst people felt this research was a good thing, with a few even thinking that anyone with a rare disease had a moral obligation to take part,

there was concern about information getting into the press and public domain, which could result in stigma or exclusion.

4. Data linkage

4.1 Awareness of data linkage

There was some awareness of specific examples of data linkage (e.g. awareness that various data were brought together to enable online purchase of car tax) but low awareness of it as a named concept or trend. Respondents were familiar with commercial organisations sharing data (e.g. Facebook), and regarded targeted advertising as both a benefit of data linkage, but also a potential nuisance. The DWP identifying benefit cheats was spontaneously mentioned as a beneficial use of data linkage, although it was felt that more could be done to catch those who flouted the system (e.g. linking Facebook data and benefit payments). Unsolicited calls about PPI were also thought to be the result of data sharing rather than opportunistic cold calling.

4.2 Health data linkage

Other than some sharing of medical records (i.e. within the NHS), data linking involving health data was not something people had given much, if any, thought to. When prompted, there was a perception from some that more linkage could usefully be done in, for example, social and mental health services to provide a more joined up, efficient administrative service. But how else health data might be 'linked' was not easily spontaneously envisaged.

A key finding from exploring the examples of data linkage was that health data linking seems more acceptable when at aggregate level rather than individual level; at an individual level there was the chance of the individual being blamed or 'told off' for something.

The notion of the government linking data raised some cynicism; fears were expressed about the government taking something away from people, especially in this economic climate.

4.3 Examples of data linkage

Facebook user data – GPS locations/use of apps, etc - being potentially linked and used for targeted advertising

The increasing powers of Facebook were acknowledged, and there was some acceptance of the lack of choice and control for Facebook users, and also users' friends who may be involved without their permission (e.g. in users' photos). In general there was resignation towards targeted advertising, although data mining including GPS location and how that data could be used, and by whom, was greater cause for concern; the idea of it being known 'where you are' at any time is an uncomfortable one.

Anonymous loyalty card data being used to influence/gauge public health programmes AND Loyalty card data being linked with personal patient health records

The first part of this linkage example was found broadly acceptable, because anonymous. But the second part sparked lively debate and strong reactions. The idea of the government snooping on people's purchasing behaviour in order for their GP to tell them later they are buying the 'wrong' food or eating things that are bad for them was considered to be outrageous! Supermarkets were perceived as hard-nosed commercial organisations, who aggressively promote unhealthy products in any case, and would not care if people are obese or have dementia. Another significant objection related to the integrity of the data (e.g. loyalty card data could not determine if an individual was purchasing for their own individual consumption or not).

Anonymous household energy use being linked to cold weather payments and hospital admissions

People did not understand the timing and mechanics of this linkage and this resulted in cynicism that the Government would try to reduce payments in some way.

***Pre-term birth data being linked with educational achievement* (shown in London groups)**

A minority saw a potential benefit if pre-term children are indeed found to have extra learning needs, but the majority responded negatively. There was fear of stigma (e.g. when applying to primary school) and C2DE respondents feared blame (e.g. for smoking or drinking during pregnancy).

***Medical research charity and/or pharmaceutical company linking mortality data with health records and statin use* (shown in London groups):**

No objections were raised to the principle of linking these data, but respondents tended to feel it should be the pharmaceutical company researching its own drugs.

Routine Government administrative data linking education records, criminal records, DWP data and income data – e.g. British Household Survey

The purpose of this data collection seemed relatively vague; some welcomed the idea that those playing the system might be identified, but others felt a sense of Big Brother collecting data that might disfavour people financially or incriminate them.

4.4 Communications about data linkage

The key fears and scepticisms to address in a positive communication of data linkage to the public are: what's the point? is it a waste of resources? how might I and others benefit? what might leak out that could be harmful to someone? what care is going to be taken with the data? what money interests might be involved? Clarity, transparency and reassurance are therefore required.

D. RESEARCH CONCUSIONS

1. General public's views on different kinds of personal data

    1.1 With data recognised as a valuable commodity used for both legitimate and illegitimate purposes, the amount of data people release about themselves is a salient issue for most people today. Fear of personal data falling into the 'wrong hands' is common.

    1.2 People acknowledge that their personal data are often held by others such as public officials and customer service personnel, and while it is recognised that this might be inevitable, it still leaves people with a sense of powerlessness that they do not control their own data.

    1.3 Overall there is fairly widespread wariness about being watched / snooped on by the government, corporations and criminals.

    1.4 However, in this age of information and technology, releasing personal data about ourselves is recognised as inevitable if we want to participate fully in society and contemporary life. It is an extreme choice to opt out.

    1.5 The main risks associated with the collecting and storing of personal data are being a victim of crime and loss/theft, nuisance marketing, and a general sense of Big Brother watching.

    1.6 The main benefits associated with the collecting and storing of personal data are convenience, advantageous offers and more efficient customer service, and possibly some sense of belonging.

    1.7 Younger people (Facebook generation) seemed to be more accepting/less questioning than older people around the giving/tracking of data. Older people may either be more wary/worried because they have seen changes/advances over time, or less aware overall of what is going on.

    1.8 The reason the pro-privacy sample tend to be actively cautious is typically due to some particular knowledge/experience (theirs or others').

2. General public's views on health data

    2.1 Health/medical data are perceived to some extent to be different from other kinds of data, primarily because there is a perceived **unquestionable benefit** to people in terms of experts having information about their health, in relation to illness or avoiding it.

2.2 The personal NHS record is top of mind in terms of health data; this is regarded as sensitive because personal and private, yet acceptable (wanted) to be shared within a medical context. There is a relatively high level of trust around data security within the NHS, at least as it currently stands.

2.3 Aggregated health data is less top of mind, but also felt to be of clear benefit – to society and individuals within it, both now and in the future.

2.4 There are superficially no/very few objections to medical data being used for the 'general good' (perceived as helping find cures and causes), provided commercial gain is not the priority.

2.5 Our research suggests the public may sometimes have a simplistic view of what is involved, in relation to health data collection, and be unaware of complex ethical considerations.

2.6 Overall, within medical research and monitoring, anonymity/consent issues are paramount, with fears centring on individuals being identified and stigmatised or disadvantaged (in their daily lives), or exposed via the media in some way.

3. Perceptions when different kinds of personal data are linked

3.1 There is some awareness of data linking in the context of government departments (e.g. car tax) and commercially (one company to another); but overall 'data linkage' is conceptually more complex for people to get their heads around.

3.2 Some see immediate benefit in the realm of catching out fraudsters etc; but there is also cynicism that when the 'government' is involved, it will be looking to pay less or do less.

3.3 If commercial organisations are involved, members of the public feel they need to be on their toes to see if there is true gain to them, or just to those selling to them.

3.4 Overall, when attention is drawn to data linking, there is some unease around both the purpose, and complexities, with the sense of Big Brother potentially prevailing.

3.5 The examples of data linking discussed in the research provoked much debate.

3.6 Any linking resulting in the **individual being targeted with specific messages** prompts discomfort and resistance. The expectation is that blame and desired behaviour change will be implicit, which fundamentally threatens the concept of free will. The lower socio-economic classes can feel particularly defensive.

3.7 Linking data (health with other) on an aggregate/anonymised level is viewed differently. This can potentially be perceived as for the greater good of society (and thus individuals within it). The arguments here are more around the integrity/accuracy of the data.

3.8 This research suggests the public does not seem particularly sensitive about who conducts research involving linking of health data, providing the objective is to increase knowledge around the causes and cures of ill health.

We are a global charitable foundation dedicated
to achieving extraordinary improvements in
human and animal health. We support the
brightest minds in biomedical research and the
medical humanities. Our breadth of support
includes public engagement, education and the
application of research to improve health.
We are independent of both political and
commercial interests.