Public Services and Government

The Scottish Government

Public Acceptability of Cross-Sectoral Data Linkage

Deliberatative Research Findings

r social research

PUBLIC ACCEPTABILITY OF CROSS SECTORAL DATA LINKAGE: DELIBERATIVE RESEARCH FINDINGS

Dr Sara Davidson and Christopher McLean, Ipsos MORI Scotland, with Professor Sarah Cunningham-Burley and Dr Claudia Pagliari, Centre for Population Health Sciences, University of Edinburgh

The views expressed in this report are those of the researcher and do not necessarily represent those of the Scottish Government or Scottish Ministers.

Table of Contents

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

Introduction

The Scottish Government is working with a wide range of partners to establish a collaborative strategic framework that will facilitate cross-sectoral data linkages for research and statistical purposes. In March 2012, it published the consultation document, *A Scotland-wide Data Linkage Framework for Statistics and Research* (Scottish Government 2012), which sets out its proposed strategy and a set of 'Guiding Principles' for linkage.

In parallel with, and to supplement the findings of, the consultation, the Government had a number of meetings and discussions with stakeholders. It also commissioned Ipsos MORI Scotland, along with Professor Sarah Cunningham-Burley and Dr Claudia Pagliari from the Centre for Population Health Sciences at the University of Edinburgh, to undertake a series of public deliberative events. This report presents the findings of the events.

The overall aim of the events was to explore the views of the public on the acceptability of linking personal data for statistical and research purposes, thereby identifying particular sensitivities and potential barriers to public confidence and exploring mechanisms for overcoming concerns.

Research methodology

The study was conducted using deliberative methods in recognition of the complexity and potential unfamiliarity of the topic, and thus the need for participants to be appropriately informed in order to meaningfully consider the relevant issues.

Three half day workshops were held; in Stirling, Inverness and Glasgow, between 26 May and 9 June 2012, with participants recruited to be broadly representative of the Scottish population.

Thirty participants were recruited for each workshop with the aim of ensuring that around 25 attended on the day. In the event, twenty-four attended the Stirling workshop, 22 attended the Inverness workshop and 27 the Glasgow workshop. Attendees were representative of the wider pool of recruits.

Key findings

General attitudes to organisations holding and using data about individuals

When participants were asked how they felt about organisations holding information about them, they tended to begin by expressing concerns about an encroaching " big brother" or "surveillance society" and/or about the amount of data on individuals that is collected and used in the commercial sphere.

The term "big brother" society was used to refer generally to the large amount of data that is collected on individuals (across both the public and commercial

spheres), and also to a proliferation of surveillance mechanisms such as CCTV, electronic tagging and mobile phone tracking. The overall feeling was that the monitoring and recording of different aspects of people's lives has gone "a bit far", compromising privacy.

Most participants were acutely aware and strongly critical of the tendency for commercial actors to sell individuals' details to each other for use in targeted marketing campaigns. Several also expressed concern about the growing ease with which such actors can "profile" individuals by drawing on data from multiple and diverse sources, including social networking sites like Facebook. There was a perception that this provides fertile ground for scams, fraud and identity theft.

Spontaneous detailed comments about the holding of personal information by public bodies were sometimes slower to emerge – or at least less well formed – than those relating to commercial actors. Nevertheless, when prompted for their views, virtually all participants engaged keenly with the subject.

Many said that they generally trusted public bodies more than commercial organisations with their personal data. The NHS tended to be seen as particularly trustworthy due to the fact that health professionals are expected to abide by a moral code of conduct as part of their job and, more generally, to serve or help the public.

Still, among a significant minority of participants, there was scepticism around the extent to which public bodies could be trusted to look after data and use it appropriately. This was in part fuelled by high profile cases of data losses and data breaches, and a perception that some public bodies are active in selling data to commercial organisations.

Many participants were also conscious of the fallibility of information technology and the difficulty of creating entirely "fool proof" security systems and procedures. There was particular concern about the potential for hacking, which tended to reflect personal experiences of fraud and other scams in the commercial domain, as well as media stories of security breaches in large organisations, such as LinkedIn[1]. Other participants spoke less about the fallibility of systems and more about security risks arising from the "human factor" in public bodies; specifically, the potential for "human error" in data handling and for officials to behave indiscreetly, unscrupulously or corruptly.

While most participants acknowledged that public bodies need to hold data on individuals, some contended that they do not always seem to make effective use of the information they have, in contrast, others felt that decision makers can place too great an emphasis on numerical data and 'statistics', which can lead to the crude categorisation of individuals or groups and result in labelling, stigmatisation and discriminatory treatment. It was also felt that reliance on descriptive statistics could lead to policies and spending plans that fail to reflect the myriad of ways in with social and other problems are experienced across the population.

---

[1] On 6th June, three days prior to the Glasgow event, Linkedin was in the news following a security breach with resulted in around 7 million of its users' passwords being posted online.

Unprompted views around cross-sectoral data linkage

Unprompted views around cross-sectoral data linkage were sought prior to an informational presentation about the proposed Data Linkage Framework.

While virtually all participants recognised potential benefits of data linkage, most also had questions and concerns about it. It was commonly felt that linkage could lead to increased negative "labelling" of people. More specifically, there was concern about the potential for labels to carry across sectoral boundaries and result in individuals or groups experiencing discriminatory treatment or stigma in multiple spheres.

Concern was similarly expressed over the possibility of linked data being used for commercial or political purposes. There was a strong consensus that such uses would not be acceptable.

A number of participants contended that linkage would increase the likelihood of data security breaches, both because more people would have access to more data, and hackers would be able to obtain a significant amount of information about individuals "in one hit". Often these concerns were based on a mistaken assumption that linkage would result in the creation of one super-database of information that would be 'warehoused' for use by multiple organisations.

Such concerns were also often premised on an assumption that the data would include individuals' names and other personal information. When participants were reminded that this would generally not be the case, many immediately became more comfortable with, or indeed positively disposed towards, the idea of linkage.

However, a small number of participants were keen to emphasise that, even with anonymisation, there would be the potential for groups to be negatively profiled and labelled. Other, particularly IT literate, participants contended that anonymised data could always be linked back to personal identifiers, by anyone with the necessary knowhow.

The Data Linkage Framework

Participants broadly supported the overarching objectives of the Data Linkage Framework. However, they had specific concerns on which they sought reassurance, which tended to centre around the questions of:

- who would oversee the operation of the Framework
- who would have access to linked data, and specifically whether this would include commercial companies
- how individuals' privacy would be protected
- how the data would be kept secure
- where overall accountability would lie if linked data was lost or stolen

These concerns were discussed in relation to data linkage generally and there was little explicit differentiation between particular sector-to-sector linkages, despite prompting by the group facilitators.

Perceptions of the draft Guiding Principles were somewhat mixed. On one hand, there was a view that the Principles go some way to addressing the main areas of concern and provide reassurance that data linkage will be carried out appropriately and securely. On the other hand, the Principles were commonly considered to be too "vague" and therefore open to interpretation and manipulation by vested interests.

Participants identified a number of safeguards that could be implemented to maximise public confidence in the Framework. These included:

- a requirement that anyone applying to use linked data must provide a strong justification to a commission or panel as to why their research is in the public interest
- publishing all processes and procedures surrounding data linkage, as well as details of who is undertaking research using linked data
- establishing an oversight body, comprising highly qualified professionals and, potentially, lay members, with responsibility for granting or refusing data linkage requests, ensuring that the Principles are upheld, and administering sanctions
- establishing accountability by placing data linkage under ministerial remit or the auspices of an independent professional or senior civil servant
- ensuring that explicit consent is obtained for uses of data containing names or other direct identifiers and that, in the process, clear parameters are set around what is being consented to (i.e. the type of research)
- requiring that consent for the uses of data containing names or other identifiers be obtained from data subjects themselves or from their next of kin, and preventing any oversight body from granting proxy consent
- ensuring all electronic systems used by individuals and organisations with access to linked data meet a minimum security requirement that is reviewed and updated frequently
- ensuring all researchers and officials with access to linked data are appropriately vetted through mechanisms such as: a certified training course; an accreditation scheme; or an assessment scheme similar to Disclosure Scotland.
- imposing strict sanctions on individuals and/or organisations responsible for any breaches of the Principles and specifying the range of possible sanctions within the Principles

There was broad support for the objectives of *Beyond 2011[2]*, which did not raise any new privacy issues. Generally 10 years was considered too long a gap between censuses and several participants questioned whether the census represents value for money given that the data soon becomes obsolete.

---

[2] A project being run by the National Records of Scotland to assess alternative options for producing population and socio-demographic statistics, including the use of administrative data sources.

There was a strong appetite for ongoing public engagement in the development of the Framework and in particular for media advertising campaigns; the distribution of informational leaflets; and the establishment of a dedicated website that could serve as a 'one stop shop' for everything members of the public might want to know about data linkage.

1   BACKGROUND

1.1   Scotland is recognised as an international leader in data linkage for the purposes of research and service analytics, particularly in the health sector where robust mechanisms exist for linking various datasets holding medical records and related sources, such as the Scottish Health Survey. This has helped to generate insights into patterns of health and illness in the population, as well as to examine the impacts of new treatments or policies (see Morris *et al*,1997; Bhopal *et al*, 2011). Data linkage has also been used successfully in other sectors, such as to demonstrate the impacts of social care on children's education and future crime (see the Edinburgh Study of Youth Transitions and Crime). At present the technical and regulatory mechanisms for undertaking record linkage studies are piecemeal and cumbersome and the Scottish Government has recognised the scope for the national research environment and strategic capabilities to be enhanced by facilitating linkage across different public sector databases. It is important to emphasise that this is different from the sharing of information across agencies for the purpose of tailoring services (e.g. clinical care) to individual needs or to identify individuals at risk.

1.2   The Scottish Government is working with a wide range of partners to establish a collaborative strategic framework that will enable cross-sectoral data linkages for research and statistical purposes to be conducted safely, securely, legally, ethically and efficiently. In March 2012, it published the consultation document, *A Scotland-wide Data Linkage Framework for Statistics and Research* (Scottish Government 2012), which sets out its proposed strategy and a set of 'Guiding Principles' for linkage. The strategy is premised on the principle of protecting individuals' rights to privacy whilst maximising the potential of data linkage for research and statistical purposes.

1.3   A key barrier to further linkage identified in the consultation document is concern among data custodians around the public acceptability of the process. Along with uncertainty around the legalities of linkage, this has resulted in custodians at times erring on the side of caution and refusing data linkage requests. Understanding public acceptability and developing safeguards to address any concerns is therefore essential if the benefits of cross-sector data linkage are to be fully realised.

1.4   While research that has focused specifically on the acceptability of cross-sector data linkage is quite limited, there is a reasonable body of scholarship on public attitudes towards linking data for health research, reasons for participating in health research and different approaches to gaining consent.

1.5   Jones and Elias (2006) have suggested that public perception in the UK is itself a barrier to effective data linkage in the UK context. However, although concerns have been identified in research on public attitudes, such studies have not found outright rejection of the aims and objectives of data linkage for research, nor of the methods through which these might be achieved (Aitken 2011). Existing evidence regarding health-related data suggests there is general support for the use of personal medical data for research (Haddow *et*

al. 2007). Where surveys have asked for consent for data linkage to the survey data collected, response rates have been good (e.g. the Growing Up in Scotland study), again suggesting a degree of support.

1.6   Nonetheless, concerns about privacy and confidentiality are raised in different research contexts and the tension that exists between these and the wider public interest (see for example Heath 2010 and Willison *et al.* 2007). Several studies have suggested that the purpose of the research plays a role in shaping public acceptability, with wide social benefit considered important, such as improvement in services (see for example Willison *et al.* 2007 and Aitken 2011). Who has access to data is also an important influence and concern, and there seems to be a degree of public unease about commercial involvement (Trinidad *et al* 2010). Damschroder *et al* (2007), amongst others, have found that trust is key to whether or not people are willing to share data from their medical records. Further, components of trust appear to relate to systems issues, such as security, research relevance (for public/patient benefit) and trust in researchers (see for example, Hunter *et al,* 2009).

1.7   The Scottish Government's consultation document and the Data Linkage Framework set out therein sought to address some of these public acceptability issues, whilst soliciting views on key questions pertaining to the ongoing development of the Framework.

1.8   In parallel with, and to supplement the findings of, the written consultation, the Government had a number of meetings and discussions with key stakeholders. It also commissioned Ipsos MORI Scotland, along with Professor Sarah Cunningham-Burley and Dr Claudia Pagliari from the Centre for Population Health Sciences at the University of Edinburgh, to undertake a series of public deliberative workshops in order to provide an enhanced evidence base on the public acceptability of the proposed Framework. This report presents the findings of these events. Responses to the written consultation can be viewed on the [Scottish Government's website](#).[3]

Research objectives

1.9   The overall aim of the research project was to explore the views of the public on the acceptability of linking personal data for statistical and research purposes, thereby identifying particular sensitivities and potential barriers to public confidence and exploring mechanisms for overcoming concerns.

1.10   Within this overarching aim, the more specific objectives were to:

- identify particular concerns or sensitivities amongst members of the public around the sharing and linking of data within and between sectors
- identify whether any particular sector-to-sector linkages raise levels of concern about privacy
- test the extent to which the draft 'Guiding Principles' reassure participants that data linkage will be governed appropriately

---

[3] The consultation analysis report is available here:
http://www.scotland.gov.uk/consultationanalysisdatalinkage

- identify what safeguards could be put in place to maximise public confidence
- investigate the extent to which the public support the objectives of the *'Beyond 2011'* project and the extent to which these raise further privacy concerns
- investigate whether the public have views about on-going public involvement and how this might be achieved.

Structure of the report

1.11   The next chapter sets out the methodology that was adopted for the research. Chapter 3 begins by examining participants' general attitudes towards organisations holding and using information about them, before considering their unprompted reactions to the idea of data linkage. Chapter 4 explores in detail participants' reactions to an informational presentation on the proposed Data Linkage Framework, including the extent to which their initial attitudes to data linkage changed in light of the presentation. Finally, Chapter 5 summarises the key findings from the research and the implications of these for the ongoing development of the *Scotland-wide Data Linkage Framework for Statistics and Research.*

2 RESEARCH METHODOLOGY

2.1   The study was conducted using deliberative techniques; a form of qualitative research. Rather than attempting to identify pre-existing attitudes to data linkage; a concept which may not have been immediately relevant to participants nor fully understood; deliberation allowed for the sharing of information and expertise, the development of considered responses and the exploration of possible strategies for a socially acceptable Framework.

2.2   Three half-day deliberative workshops were held; in Stirling, Inverness and Glasgow between 26 May and 9 June 2012[4].

Recruitment

2.3   Participants were recruited face-to-face in their homes between 14 May and 7 June 2012, using a questionnaire specially designed for this purpose.

2.4   To ensure that a broad range of people were engaged in the research, quotas were set on: sex; age; working status; socio-economic grade; ethnicity; disability; parental status; and area. Further, participants with varying levels of trust in public bodies were recruited[5] as previous research has found trust to be important in shaping attitudes to data linkage for health research (see Aitken 2011).

2.5   Individuals who worked in market research, media, advertising or PR; and those who had attended a group discussion or workshop in the previous 12 months were excluded from the research.

2.6   Thirty people were recruited for each workshop, with the aim of ensuring that around 25 attended on the day. In the event, and as Table 5.1 shows, 24 people attended the Stirling workshop, 22 attended the Inverness workshop and 27 the Glasgow workshop. In each case, attendees were representative of the broader pool of recruits (Appendix A provides a full breakdown of the profile of attendees). All attendees received £40 as a 'thank you' for their time and to cover any expenses incurred.

Table 5.1: Number of participants attending each workshop

Location	Number of participants
Stirling	24
Inverness	22
Glasgow	27

---

[4] All of the workshops were undertaken prior to the news coverage of the loss of a Glasgow City Council laptop containing personal data of customers and businesses (this incident first appeared in the mainstream media on the 11 June). The views expressed by participants in the workshops were therefore unlikely to have been affected by this data loss incident.
[5] Participants level of trust in public bodies was established using the following question, included in the recruitment questionnaire: '*I will read you a list of different types of people. For each, would you tell me if you generally trust them to tell the truth, or not?*' The Scottish Government; local councils; researchers in universities; the NHS; The Police.

Structure of the workshops

2.7 The workshops comprised a mix of plenary and small group sessions. For the latter, participants were divided into three groups on the basis of their age (18 to 34 years, 35 to 49 years and 50 years and over) to allow for the identification of any variation in views by life stage.

2.8 Table 5.2 shows the structure of the workshops and summarises the purpose of each stage.

Table 5.2: Structure of the workshops

Element	Purpose/content
Initial plenary session	To welcome participants and to provide an outline of the scope of the study – i.e. data linkage for research and statistical purposes and not for the sharing of information about individuals to inform service delivery – and the key concepts that would be used throughout the day.
Breakout groups (1)	Initial warm-up discussions to explore general attitudes towards data collection and use, and to gauge initial reactions to the concept of data linkage.
Presentation on the proposed Data Linkage Framework and Q&A	The presentation covered detailed information on: • what data linkage is and how it is done • benefits and risks of data linkage (with reference to case study examples) • why a Data Linkage Framework is necessary • the proposed Framework and Guiding Principles.
Lunch break	
Breakout groups (2)	To explore participants' perceptions of the proposed Data Linkage Framework, including: • their overall reactions and more specific reactions regarding particular sector-to-sector linkages • the extent to which they felt that the Guiding Principles provide appropriate governance frameworks • their suggestions for safeguards to address concerns and maximise public confidence • their views around the Beyond 2011 project.
Plenary session	A summing up of the key messages from the event and completion of a post-workshop questionnaire.

Workshop materials

2.9 All the materials used in the workshops were designed by the researchers at the University of Edinburgh and Ipsos MORI Scotland with input from the Scottish Government. These included a topic guide which was used by moderators to facilitate the break-out discussions, a PowerPoint presentation with detailed information on the proposed Data Linkage Framework and a post-workshop questionnaire (copies of all materials are provided in Appendix B).

Analysis

2.10 All discussions that took place at the workshops were recorded and then transcribed with the consent of participants. Group facilitators also compiled summary field notes at the end of each event. In addition to this, the project

team held regular meetings, throughout the fieldwork process, in order to discuss emerging findings.

2.11   At the end of the fieldwork, the researchers conducted a brainstorming session to identify the top-level findings and implications. This culminated in the creation of a code frame of substantive themes and sub-themes. Transcripts were then systematically analysed for key points and illustrative verbatim comments. Any new sub-themes which emerged at this stage were integrated into the code frame.

2.12   This method ensured that analysis and reporting of the data was rigorous, balanced and accurate, and that key messages were brought out. It was also flexible enough to allow for links and connections across different pieces of data to be made, and for moments of interpretive insight and inspiration to be recorded.

Interpretation of qualitative findings

2.13   The findings presented in this report were derived using qualitative methods. Unlike large surveys, qualitative social research does not aim to produce a quantifiable or generalizable summary of population attitudes, but to develop a deeper understanding of the range of issues influencing attitudes as well as identifying key attitudinal tendencies that are likely to be prevalent across society. Qualitative research is particularly useful when exploring complex or hard-to-understand areas, such as cross-sectoral data linkage, where it can be difficult to get a true sense of public attitudes from questionnaires. The integration of 'deliberative' approaches aids this process, since participants are given the opportunity to consider their feelings towards the concept both before and after it is explained to them. This insight often influences attitudes and reveals a more nuanced and informed set of considerations, which can be useful for informing policy making.

3 GENERAL ATTITUDES TO ORGANISATIONS HOLDING AND USING DATA ABOUT INDIVIDUALS

3.1 This chapter explores participants' general attitudes towards organisations (broadly defined) holding and using information about them, and the main considerations that subsequently underpinned their views in respect of data linkage. The latter part of the chapter explores initial, unprompted reactions to the idea of linkage, and related key discussion points.

3.2 When participants were asked how they felt about organisations holding information about them, they tended to begin by expressing concerns about an encroaching " big brother" or "surveillance society" and/or about the amount of data on individuals that is collected and used in the commercial sphere. More detailed comments relating specifically to the holding of information by public bodies tended to come later but, nonetheless, were inextricably bound with these initial themes.

Big brother society

3.3 The term " big brother" society was used to refer generally to the large amount of data that is collected on individuals (across both the public and commercial spheres), and also to a proliferation of surveillance mechanisms such as CCTV, electronic tagging and mobile phone tracking. While a minority of participants commented that they didn't mind or, indeed, that they welcomed, increased surveillance – either because it made them feel safer or because they had "nothing to hide" – the overall feeling was that the general monitoring and recording of different aspects of people's lives has gone "a bit far", compromising privacy.

> Everybody has got [a mobile phone], they [can] be tracked anywhere […] If I've done something wrong they'll find me. As soon as I start using [it]. They can now put into your phone a wee bug thing, you don't know it's there, they know what part of the street you're standing in.
> (Male, aged 50 or over, Stirling)

> If you're selling your house, for instance, everyone on the internet can see round your house and know what you're getting for it when it's sold. I think that's a right breach of privacy and shouldn't be allowed.
> (Female, aged 50 or over, Inverness)

3.4 Nevertheless, as the discussions progressed it became clear that even the staunchest critics of the big brother society were in fact fundamentally ambivalent on the matter, opposing data collection and surveillance that impacted on their own freedoms or privacy, but supporting that which they saw as benefitting either themselves or their communities (measures to monitor the movements of criminals, paedophiles and terrorists received particular mention in this regard). The following excerpt from a discussion between young people at the Stirling event is illustrative:

Male 1: [In] a few years to come they are actually going to chip us and it's for our benefit right and it's so we have ID and don't need a passport because all they are going to do is scan this wee thing that is attached to your skin. That is basically so they know where you [are at] all times. It's just like what you do with your dog… Don't get me wrong, there should be folk [who are chipped]; paedophiles and that…

Male 2: Paedophiles and rapist and that should be chipped so the police can go: "Oh, right, there he is in that street at such and such a time".

Female 1: But that will benefit everybody in society.

3.5 The contingent nature of views surrounding the privacy versus public interest dichotomy was a recurring theme across the events and is returned to over subsequent sections.

Data in the commercial domain

3.6 Participants spoke extensively and vociferously about the amount of personal data circulating in the commercial domain, typically bemoaning the high number of unsolicited sales calls they receive from companies which appear to have obtained information on their circumstances. Most participants were acutely aware and strongly critical of the tendency for commercial actors to sell individuals' details to each other for use in targeted marketing campaigns. Several also expressed concern about the growing ease with which such actors can "profile" individuals by drawing on data from multiple and diverse sources, including social networking sites like Facebook. There was a perception that this practice provides fertile ground for scams, fraud and identity theft. Indeed, a significant proportion of participants had experienced bank fraud and this led many of them to reflect on the apparent ease with which such offences are committed.

Advertising companies etc. will take information from your PC or what you've been looking [at] for your internet searches etc., and they are selling that in to advertising folk so that when you go onto websites and things it will come up with things that you are more likely to buy. They can basically run a personality check on me.
(Male, aged 18-34, Stirling)

Now everybody and their uncle with a cheap computer and a mobile phone can sit outside your house and get all your information. Every penny they took [from my bank account].
(Male, aged 50 or over Inverness)

3.7 It was clear that negative experiences of data collection or use by commercial actors had left many participants feeling wary of *any* organisations having their personal information, which impacted on subsequent discussions around data linkage.

Public bodies

3.8 As already mentioned, spontaneous detailed comments about the holding of personal information by public bodies were sometimes slower to emerge – or at least less well formed – than those relating to commercial actors. Indeed, some participants stated that they had "never really thought about" this matter or that it is "just part of the normal running of things" or "a fact of life" to which they are resigned. Nevertheless, when prompted for their views, virtually all participants engaged keenly with the subject. Their comments centred around three inter-related themes: trust; data security; and data use.

Trust

3.9 Many participants said that they generally trusted public bodies more than commercial organisations with their personal data. As Aitken (2011b) found, the NHS tended to be seen as particularly trustworthy due to the fact that health professionals are expected to abide by a moral code of conduct as part of their job and, more generally, to serve or help the public.

> If you give your information to the Government etcetera, you kind of feel a bit more trusting that it is going to be used for a valid purpose.
> (Male, aged 35-49, Glasgow)

> I think you just assume that all your information is kept private. You just assume because [doctors] are obviously professionals; that's kind of their job description.
> (Female, aged 18-34, Stirling)

3.10 Still, a significant minority of participants held considerably more sceptical views and there were four main reasons for this. The first was a wider lack of trust in public officials to act in the public interest, rather than for personal or political gain. Discussing this point, participants tended to refer to high profile scandals (for example, the MPs expenses row and the News International affair) and also to more localised cases of officials behaving corruptly or appearing to withhold information from the public – In Stirling there was specific reference to perceived attempts by the NHS to conceal or play down hospital infections. However, and as is discussed more fully later in this chapter, it was apparent that generalised distrust in public bodies also stemmed from feelings of disempowerment or disadvantage at the hands of 'the system' – for example, in respect of service provision, as well as broader processes such as 'labelling'.

> If you have been a trouble maker as a child, then you are automatically a trouble maker when you're older; the same with the Council too – if you've got a nice enough face or whatever, then you do get stuff; you get a nice house with a big garden and the person they don't like is in the corner in a big high flat with five bairns and that's it.
> (Female, aged 18-34, Stirling)

3.11   Second, there was repeated reference to high profile cases of public officials inadvertently losing personal data by leaving laptops and data sticks on trains or in bins or skips. These references often formed part of wider discussions about the potential for "human error" and "rogue" behaviour in all organisations, and the implications of this for data security; themes which are also returned to in the next section.

3.12   Third, among a significant proportion of participants, it was simply assumed that public bodies are actively selling individuals' data to commercial actors. This assumption was in part based on stories participants had come across in the media but, more commonly, on personal experiences. For example, several people recounted occasions on which they had been contacted by a commercial organisation in relation to a (sometimes recently diagnosed) health problem, contending that the only way the organisation could have known about their condition was through their GP or another branch of the NHS. Similarly, a couple of participants at the Stirling event made vague references to hearsay about the Council "selling off information to companies". There was suggestion that such behaviour by public bodies is symptomatic of the fact that they are increasingly short of funding and so under pressure to find additional sources of revenue.

3.13   A few of the more trusting participants expressed doubt that public bodies would sell personal information, pointing out that the Data Protection Act and professional codes of practice ward against this. However, their comments tended to hold little sway among their sceptical peers.

I got phone calls from a place in Dunblane saying: "Would you like to come and get treatment on your back?" How did they know I had problems with my back?
(Male, aged 50 or over, Stirling)

I think a lot of it is down to budget constraints. They will sell that information because their particular department needs the money. [T]he big companies, the drug companies and all the rest of it know this; they know all these government departments and things are strapped for money so they will offer money.
(Male, aged 50 or over, Inverness)

3.14   Some of the older participants alluded to a more general blurring of the boundary between the civic and commercial spheres owing to the contracting out or gradual "privatisation" of public services or elements thereof (the NHS tended to be the main focus here but in Glasgow there was also reference to the creation of Arms Length Organisations, or 'ALEOs', to run council services); a development that was seen as heralding a shift towards decision making based on financial, rather than public interest, imperatives, and, consequently, to have negative implications for the security and use of personal data.

> The hospitals and that they're [be]coming more private all the time. It's selling the information; drugs and scans. There is no privacy.
> (Male, aged 50 or over, Stirling)

Security

3.15 Consistent with Aitken's (2011b) findings, many participants were all too conscious of the fallibility of information technology and the difficulty of creating entirely "fool proof" security systems and procedures. There was particular concern about the potential for hacking, which tended to reflect personal experiences of fraud and other scams in the commercial domain, as well as media stories of security breaches in the largest and (ergo in participants' minds) best protected of organisations – there were specific references to the FBI, CIA and companies such as LinkedIn[6].

> Is anything secure? Some of the biggest data companies have lost millions and millions of people's data, so I just don't believe that anything is secure.
> (Male, aged, 50 or over Glasgow)

3.16 As already indicated, other participants spoke less about the fallibility of systems and more about security risks arising from the "human factor" in public bodies; specifically, the potential for "human error" in data handling and for officials to behave indiscreetly, unscrupulously or corruptly – whether this be in terms of sharing or selling data, or using it for "covert" purposes.

> I think I actually trust [public bodies] quite heavily, but there is always individuals; individuals will lose a memory stick, which unfortunately will [hold] quite a lot of useful data.
> (Female, aged 35-49, Inverness)

> It's individual people. It's people's nature that they like to know others' business…and they like to spread gossip, they like to cause mischief.
> (Male, aged 50 or over Inverness)

3.17 As is discussed more fully towards the end of this chapter, such security-related concerns were often augmented when the discussion turned to the subject of data-linkage.

The use of data

3.18 Across the events, most participants acknowledged that public bodies need to hold data on individuals. While this point was most commonly made with reference to the importance of doctors and hospitals holding comprehensive patient records, the use of data to better design and target services was also recognised.

---

[6] On 6th June, three days prior to the Glasgow event, Linkedin was in the news following a security breach with resulted in around 7 million of its users' passwords being posted online.

I think it's just an essential, whether you like it or not, in modern life. If […] there is no school for your kids to go [to] because they didn't have the statistics to say they needed a school there, then you wouldn't be too happy. It's just a fact of life; they need [data].
(Female, aged 35-49, Inverness)

3.19   At the same time, however, some participants contended that public bodies do not always seem to make use of the information they have. This view was often, though not exclusively, expressed by people who felt let down or in some way short changed by 'the system'; for example, one participant spoke incredulously about the fact that a known paedophile had been moved into her area, despite the area containing a high proportion of families with young children. Another contended that services as diverse as social work and refuse collection do not seem to be targeted appropriately, and another again commented that sections of society that are most in need of services, including the elderly, often bear the brunt of spending cuts or freezes.

3.20   Somewhat paradoxically, other participants felt that there is *too much* focus on (quantitative) research data and statistics in decision making, with the effect that individuals and groups are often crudely categorised and "labelled". As Aitken (2011) found, this was seen to result in discriminatory treatment and/or stigma, as well as policies and spending plans that fail to reflect the myriad of ways in with social and other problems are experienced across the population.

See the deprivation and that, there is people who are […] sitting there, maybe down in London, going: "wait a minute; that area there that's poor that area there… [we] better share out to the poor area" and [other] pockets get missed out. They aren't in the real world these people.
(Male, aged 35-49, Stirling)

3.21   As with concerns about data security, labelling re-emerged as a prominent theme during discussions about data linkage.

Data linkage

3.22   Unprompted views around cross-sectoral data linkage were sought towards the end of the first break-out sessions, prior to an informational presentation about the proposed Data Linkage Framework.

3.23   Participants tended to conceive of linkage primarily in terms of the sharing of information (both between and within sectors) about individuals to deliver "joined-up" services or to give agencies a better overview of individuals' circumstances. In particular, there was reference to the sharing of information between different parts of the health sector; between the health sector and social care sector; and between the social care and housing sectors. Similarly, a number of participants across the groups referred to the linking of data across agencies such as the HMRC, the DWP and housing departments/associations to identify people who

are trying to "fiddle, wangle or get round the system". Linkage for research and statistical purposes was also mentioned spontaneously, albeit less frequently.

3.24 Virtually all participants recognised potential benefits of data linkage. In relation to linkage for research and statistical purposes specifically there was a particular focus on the potential for this to deliver better intelligence about *local areas* and subsequently improved or more targeted services. Many participants, several of whom worked in the health sector or had long term health conditions, also recognised that linkage could provide evidence to support public health improvements, especially in terms of the prevention and management of chronic and terminal conditions.

> If you can work out that […] a small area of Glasgow has an increased incidence of lung cancer, and you also notice increased smoking and there's more unemployment about, you can target to address those issues. You can be more specific based on the needs of a specific area.
>
> (Female, aged 18-34, Glasgow)

3.25 Still, most participants also had questions and concerns about data linkage. These, again, tended to cluster around issues of data use and data security, but anonymity and consent were also strong themes.

Data use

3.26 People often commented that, before giving definitive opinions on cross-sectoral data linkage, they would like to know more about the ultimate objectives of linkage, who would have access to the data and, related to this, how the data might be used. This reflected the widely held view that "information can be used for good or for bad".

3.27 On the theme of "bad" uses, it was commonly felt that linkage could lead to increased negative "labelling", or as some put it, "stereotyping" of people. More specifically, there was concern about the potential for labels to carry across sectoral boundaries and result in individuals or groups experiencing discriminatory treatment or stigma in multiple spheres. Stigma on account of having a criminal record – or simply having had some involvement with the criminal justice system – received particular mention. Such concern often led participants to express a view that bodies should only be able to access data that is directly relevant to their work.

> It's not nice to think that you're having a link between housing and education, you know […] because housing is there for the parents, education is there for the child, and you can't make assumptions from a parent to a child, or the other way around: They're doing badly at school; their parents live in a council house. If you live in a council house; the kids will do badly at school. There's not a correlation between these things.
>
> (Female, aged 18-34, Glasgow)

> If you have done something 15 or 20 years ago [and] that's on your information, are you suddenly going to be subject to judgments?
> (Male, aged 35-49, Stirling)

> I think the police is a really scary one, because [there are] so many people who have been convicted and then had to go back and appeal that and different things.
> (Female, aged 35-49, Stirling)

> If you're sharing data and if [an organisation has] what they only require for their particular department or industry or whatever, and not the rest of the information then that would be fine. But if certain departments are getting access to all the information, then that would become an infringement on your rights.
> (Female, aged 35-49, Glasgow)

3.28   Concern was similarly expressed over the possibility of linked data being used for commercial or political purposes. In referring to commercial uses, participants tended to have in mind companies accessing the data for marketing purposes and "financial gain". However, there was also a perception that banks and insurance companies could use the data to 'vet' potential clients; for example, refusing a mortgage to someone who had been evicted from a council property in the past. There was a strong consensus that such uses would not be acceptable.

3.29   Concerns around possible political uses of linked data were comparatively uncommon but no less keenly felt.

> If the Government are using the details for the benefit of society, I think that's okay. But if the Government are using that data to then look at their next election campaign, or look at the independence campaign by looking at the demographics of a particular area, then I don't know if that's as acceptable. They'[d] simply be using our data for their own goals.
> (Female, aged 18-34, Glasgow)

3.30   As the above quotation helps to illustrate, a dominant perspective was that for linkage to be allowed it must hold potential benefits for the public, or at least a segment of the public, whether that be a patient group, or residents of a particular area, or type of area etc.. Notably, benefits were conceived of in fairly broad terms, from short or medium term improvements to services, to longer term medical advances or health improvements.

Security

3.31   On the issue of security, participants often reiterated concerns, outlined earlier in this chapter, about the potential for any public data to be hacked and about the "human factor" in organisations that can result in data errors, losses and misuse. A number of participants contended that

linkage would increase the likelihood of security breaches occurring, both because more people would have access to more data, and because it would be possible to obtain a significant amount of information about individuals *"in one hit"*. Often these concerns were based on a mistaken assumption that, ultimately, linkage would result in the creation of one super-database of information that would be 'warehoused' for use by multiple organisations.

> It would just make it easier for people to get that information in one hit rather than get[ting] a little bit from [here and] there, which would take longer. It makes it easier for people to steal your information.
> (Male, aged 18-34, Inverness)

Anonymity and consent

3.32   Concerns about security and about potential misuses of linked data, were also often premised on an assumption that the data would include individuals' names and other personal information. When participants were reminded that this would generally not be the case, many immediately became more comfortable with, or indeed positively disposed towards, the idea of linkage (notwithstanding their conviction that the data should not be used for commercial or political purposes).

> If it's anonymous; who cares?
> (Male, aged,35-49, Inverness)

> If it's anonymised, it's not doing any harm to anyone*.*
> (Female, aged 18-34, Glasgow)

3.33   However, a small number of participants were keen to emphasise that, even with anonymisation, there would still be the potential for *groups* to be negatively profiled and labelled – for example, one participant expressed concern that linked data could be used for eugenics and racial profiling.

3.34   Other, particularly IT literate, participants contended that anonymised data could always be linked back to personal identifiers, such as people's names, addresses and other details, by anyone with the necessary knowhow, and that this presents a major security risk requiring consideration.

3.35   Regardless of their views around the impact of anonymisation, there was a general sense in which participants felt they should be kept informed about any linkages involving their data. A number of participants took this further, arguing that data subjects should routinely be asked for their permission before their data is used.

3.36   All of the above-mentioned views and concerns surrounding data linkage are discussed further in the next chapter, which explores participants' reactions and attitudes to the Data Linkage Framework.

4 THE DATA LINKAGE FRAMEWORK

4.1   Participants were given a presentation by the research team on data linkage and the Data Linkage Framework (A copy is provided in Appendix B). This covered:

- what data linkage is and how it is undertaken
- benefits and risks of data linkage (with reference to case study examples)
- why a Data Linkage Framework is deemed necessary
- the proposed Framework, including the Guiding Principles

4.2   This chapter explores participants' overall reactions to the Data Linkage Framework and the extent to which the Guiding Principles reassured them that data linkage would be governed appropriately. It also considers views around ongoing public engagement in the development of the Framework and how this might be achieved.

Overall reactions to the Data Linkage Framework

4.3   Similar to Aitken's (2011a) findings on attitudes to the linking of social care, housing and health data, most participants' initial reaction to the Data Linkage Framework was one of qualified support: They felt that it was a good idea in principle, but that its success would depend on it not being "abused" or "manipulated", either by public bodies or individuals within those bodies. It was clear that the provision of information about linkage had gone some way towards ameliorating many of participants' initial concerns about it, particularly in respect of anonymity, sanctions and, more generally, the types of research that would be conducted using linked data.

4.4   Stated benefits of data linkage that participants found particularly compelling were:

- the more efficient use of public resources – it was acknowledged that more data linkage would help to reduce duplication in data collection and research. (One participant suggested that the Scottish Government should publish details of how much money the scheme would cost to set up and run and how much it would save)
- improvements to service provision through the better use of data to identify which services are required and where
- the formalisation of procedures to improve the protection of personal data – many participants thought that data linkage was already happening in Scotland and were reassured that it would now be subject to greater control.
- benefits for the Scottish economy through investment in research and job creation

4.5   A small minority of participants remained very sceptical about data linkage, however, regarding the Framework as the "thin end of [a] wedge" that would ultimately lead to the Government holding large databases of information about every aspect of individuals' lives – the term "big brother"

cropped up again here. Several of these participants went on to express concern that the databases could be sold to commercial companies in order to generate much needed public finance.

> And what's the next step? Are we going to be integrating this, that and the other?… I'm a bit dubious, because that's where it seemed to be steering towards, the creating of resource for somebody who may not technically have our best interests at heart.
>
> (Male, aged 35-49, Glasgow)

4.6   Regardless of whether their overall initial reactions to the Framework were generally positive or negative, most participants went on to engage critically with the information provided in the presentation, raising a number of concerns and issues for discussion. These broadly reflected the main themes that had emerged earlier in the events and centred around:

- who would oversee the operation of the Framework
- who would have access to linked data, and specifically whether this would include commercial companies
- challenges involved in keeping the data secure
- how individuals' privacy would be protected; in particular to prevent unsolicited contact from commercial companies
- the potential use of linked data for greater 'profiling' or 'labelling' of individuals and groups – there was a sense in which the presentation has served to heighten these concerns for a number of participants, some of whom were moved to describe very specific personal experiences of labelling and the negative impact this had on their lives
- where overall accountability would lie if linked data was lost or stolen

4.7   These issues, along with participants' suggestions for how they might be addressed in the context of the Framework, are considered over the remainder of this chapter, which explores in detail views around the Guiding Principles.

4.8   Notably, the issues were discussed in relation to data linkage *generally.* There was very little explicit differentiation between particular sector-to-sector linkages; notwithstanding some remaining sensitivity around linkage involving criminal justice data and the potential for this to result in individuals experiencing discrimination across multiple spheres.

The Guiding Principles

4.9   To facilitate discussion of the Guiding Principles, participants were provided with a one page summary of these; a copy of which is provided in Appendix B).

4.10   Overall perceptions of the Principles were somewhat mixed. On the one hand, participants often commented that the Principles covered all of main areas of concern about data linkage that had been expressed over the course of the events, and that nothing important appeared to be missing.

Many went on to say that they felt reassured that linkage would be carried out appropriately and securely.

4.11 On the other hand, a common concern was that the Framework in general, and the Principles specifically, were too "vague" or "general" and therefore open to interpretation and manipulation by vested interests. A few participants took this further, contending that the Principles had probably been left deliberately vague so that the Scottish Government is able to justify any linkage it might deem necessary in the future. As is discussed more fully below, such concerns were compounded when participants identified a few specific clauses within the Principles which they regarded as "loopholes" that were ripe for exploitation – for example, giving an oversight body the right to grant authorisation to use named data when it is not possible or practicable to obtain consent.

4.12 Notwithstanding these mixed views, there was a consensus that all of the Principles were essential and needed be accorded equal priority to ensure the proper governance of data linkage. For example; it was acknowledged that in order to ensure privacy, the removal of names and direct identifiers was required, which in turn required security to ensure that data could not be linked back to personal details.

If you take one [principle] out then it's gone, it collapses.
(Male, aged 50 or over, Inverness)

Public interest

4.13 In order to ensure the correct balance between the benefits of data linkage and the protection of individual privacy, participants felt that a minimum requirement of any research involving data linkage should be that it is 'in the public interest'. However, there was also considerable concern about how public interest could be defined. The term was regarded as fairly nebulous and open to varied and shifting interpretation, which in turn led many participants to feel that it could be used to justify virtually any use of linked data.

Over time, opinions of public interest can change. What is determined to be acceptable now… wouldn't have been quite acceptable [ten years ago]. So that leaves a bit of a gap as well.
(Female, aged 35-49, Glasgow)

4.14 There was particular concern about how companies and political actors might interpret 'public interest'. Reinforcing findings reported in the previous chapter, participants asserted firmly that the public interest justification should not be used by elected officials to commission research aimed at furthering their own political ends, nor by companies - nor individuals/bodies sponsored by companies - for commercial gain.

A lot of what people have got a problem with is vested interest, greed, people pursuing their own agendas. This kind of evidence based stuff is the opposite of that, because this is hard facts; it's evidence, it

speaks for itself and the problem comes when politicians start interpreting it and put their own spin on it.

(Female, aged 35-49, Stirling)

Don't get me wrong, I trust scientific research, but who is backing them and who is funding them? You can get independent scientist that are getting funded and they don't know who they are getting funded off, which could make it corrupt that they are eligible for that information.

(Male, aged 18-34, Stirling)

4.15 As indicated earlier in the report, participants tended to conceptualise the 'public interest' in terms of tangible benefits for individuals and groups such as medical advancements, and improved service provision. The consensus was that any researcher making a request to access linked data should be required to justify to a commission or panel of experts why their research is in the public interest, and that the "bar should be set quite high" to ensure the system is not abused.

Governance and transparency

4.16 Perceptions surrounding the governance and transparency principle were underpinned by the lack of trust in some public bodies, discussed in Chapter 3 and, related to this, by a view that too often important decisions are made "behind closed doors", without any public involvement.

[Data linkage] might be completely transparent but I don't think necessarily everyone will believe it is transparent.

(Female, aged 18-34, Stirling)

4.17 There was a consensus that all procedures and processes surrounding data linkage should be as "accessible" as possible to help address public concerns. Specific suggestions put forward in this regard included publishing:

- full details of the requirements that individuals or organisations applying for access to linked data must meet and the procedures the oversight body would adhere to when evaluating applications
- all requests for access to linked data, including who has been making requests, how often and why
- all results and reports based on analysis of linked data

4.18 Such steps, it was felt, would enable external groups and members of the public to monitor any activity taking place under the Framework and raise objections in the event that they felt the Principles were not being adhered to.

You make it accessible and understandable. If people were all engaged in it and all understood it and all knew what was going on then people would be more likely to track the changes and see if their information's getting used for something [it shouldn't be] and pull someone up for it.

(Female, aged 18-34, Glasgow)

4.19 Participants strongly supported the proposed establishment of an oversight body to ensure the proper governance of the Framework and facilitate transparency. It was widely felt that this body should have responsibility for granting or refusing data linkage requests, ensuring that the Principles are upheld and administering sanctions as required.

4.20 Participants frequently raised the question of who would be on the oversight body, often commenting that the Principles are too vague on this point. Asked who they would like to see on the body, they commonly suggested that senior public sector professionals should have a pivotal role, although there were different views on what this would mean in practice.

4.21 Some participants favoured the establishment of a large pool of professionals representing a range of sectors (such as health and education) from which smaller panels could be drawn on a case-by-case basis, reflecting the nature of a linkage request – it was felt that this would minimise the likelihood of individuals with "vested interests" having too much influence on decisions. Others suggested that the professionals should be data linkage experts with extensive experience and training, as well as enough understanding of each others' areas of work to be able to challenge one another to ensure a rigorous process. Still others said they would like to see a mixed body of professionals and lay members drawn from a cross-section of society (two participants cited the Children's Panels as a possible model). It was felt that the professionals would be able to consider and provide advice on technical points while lay members would represent public concerns.

It's a very responsible sounding job, so you're not looking for volunteers or anything. You're looking for high quality individuals or [a] high level of training.

(Male, aged 35-49, Glasgow)

Instead of having a fixed panel, you've got 100 people to choose from, and if you know that this person has an association with a drugs company, I'm sorry you can't be on the panel. These are completely independent people from the people putting the [submission] in to have the information.

(Male, aged 35-49, Stirling)

4.22 Accountability was regarded as crucial to ensuring that governance and transparency are properly administered. Again, however, there were differing views on how accountability should operate. Some participants

felt that data linkage should fall under ministerial remit, while others felt it should come under the auspices of an independent professional or senior civil servant to ensure impartiality.

> [The person responsible for overseeing data linkage] should be elected and if not should be having an elected official looking over them […] like a minister for information.
>
> (Female, aged 50 and over, Inverness)

> I think it would have to be some kind of independent ombudsman. It couldn't possibly happen without that because you can't manage yourself. You've got to have somebody who is completely distanced from it and impartial.
>
> (Male, aged 35-49, Glasgow)

Privacy and the removal of names and direct identifiers

4.23   As discussed in Chapter 3, many participants felt more positively disposed to the idea of data linkage after being informed that, in most cases, the data would be anonymised. After the presentation they often felt further reassured on this point, sometimes commenting that they "couldn't see any problem" with data being used in this way. However, perceptions were mixed when the use of named data was considered.

4.24   Some participants, most of whom were in the youngest age category, said they were indifferent about the use of named data, provided they were not contacted and their data was kept secure. They tended to say that they had "nothing to hide" and so had no qualms about their personal details being used in research; or that researchers working with the data wouldn't know the data subjects, meaning there would be no threat to those subjects' privacy.

> If it is going to [result in] a cure for cancer then here is all my medical stuff; have it. That's how I would be, but I don't see why there is this whole, take your name off it and then we'll take it.
>
> (Female, aged 18-34, Stirling)

4.25   However, most participants expressed concern about the possibility of being identified in research. They felt that named data was more sensitive than non-identifying data and that consequently it required greater protection (see 'security' and 'consent' below). As a result, they expressed strong support for the privacy principle and often followed this up by repeating their assertion that commercial companies should not be granted access to their contact details.

4.26   As discussed in Chapter 3, most participants were happy for their data to be used for research if it was anonymous. Still, others remained concerned that their data could always be linked back to their contact details, and that this in turn would leave it open to misuse.

[Linkage] would require each [bit of my data] having a unique identifier that is the same. That means that if you did have somebody's personal identifier then you could go round and access all their data.

(Male, aged 35-49, Glasgow)

You can trace it back to whoever it is [...] I don't have the first clue, obviously, but there is people out there that do and they are experts in that field.

(Female, aged 18-34, Inverness)

4.27   On a related point, some participants focussed in on the statement within the principle that read: any departure from [removing names and direct identifiers] must be justified and approved. The grounds on which such action would be justified, and who would be responsible for approving it were questioned. In essence, it was felt that the statement was a "loophole" that diluted the intention and rigour of the Principle, and could further leave data open to misuse.

"Any departure from that must be justified". I'm afraid I'm still worried about that […] It seems to me it's just a get-out clause.

(Male, aged 35-49, Inverness)

Consent

4.28   There was strong agreement that consent should be obtained for the use of named data. Many participants stated that too often, consent is obtained implicitly through the 'small-print' in documents or through vague statements, and that, in the case of data linkage, therefore, explicit parameters should be set around what consent is being given for (i.e. the types of research and/or the public bodies who will have access to the data), with any divergence requiring further consent.

As long as you're fully consenting and aware of what the information is being used for then I don't see a problem. It's when it starts being used for a purpose that you didn't intend it for; that's the issue.

(Male, aged 35-39, Glasgow)

Maybe there should be a field where you say 'I agree for medical research, I agree for anything for the benefit of the community but when it comes to third party companies wanting to find out my details, no way'.

(Male, aged 18-34, Stirling)

4.29   Participants were conflicted over the frequency with which their consent should be sought for the use of their personal data. After initially suggesting that consent should be obtained each time the data might be used, many soon came to acknowledge the potentially prohibitive time and cost implications of this. They subsequently suggested that their consent could be obtained on a periodic basis (e.g. annually) or via a system that would enable them to make a request to public bodies that their data not be used for research.

4.30 A particular area of concern for participants was the prospect that an oversight body would have the authority to grant consent in cases where it was not possible or practicable to obtain consent from an individual. Although participants recognised that there would be instances where it would not be possible to obtain individuals' consent, the prevailing view was that an external body should not have the right to grant proxy consent. Only a small number of participants took a different view, suggesting that the oversight body could intervene solely in cases where all other avenues have been exhausted.

It's either the person gives consent or they don't. It shouldn't be up to anybody else whatsoever.

(Female, aged 35-49, Inverness)

…maybe the person didn't have anybody else and maybe the only way you could sort it would be an oversight body […] but for me I think it would have to be a very, very, very last resort.

(Female, aged 55+, Stirling)

4.31 A minority of participants felt that consent should be extended to all uses of linked data, not just cases where named data might be used, as this would provide greater reassurance that individuals' privacy will not be compromised. However, they did acknowledge that this would be difficult to administer and may serve as a barrier to research in the event that people habitually refuse consent out of a generalised distrust in government and/or researchers.

Security

4.32 Participants were unanimous in the view that data linkage must conform to the highest standards of security but reiterated concerns that security can never be guaranteed due to the potential for incompetence (public officials losing data sticks) or maliciousness (hackers and "rogue" employees who set out to steal data).

4.33 Participants were wary of merged datasets being stored in a single location and felt that the process of sending data from one computer to another during the linking process would further increase the potential for information loss or theft. Some participants went on to express concern that, because technology changes so quickly, public bodies and other researchers accessing data may not always have the most up to date electronic security systems in place. It was suggested that a minimum security requirement should be built into the Framework, that is reviewed and updated frequently, and that anyone requesting access to data would have to prove they can meet.

4.34 Having established that security is difficult to guarantee absolutely, participants' took opposing stances in respect of the level of risk this presented. Most felt that the risk was acceptable, provided the highest possible security systems and procedures are in place. However, a

significant minority disagreed, contending that the risk represented sufficient grounds for not proceeding with data linkage.

Access and personnel

4.35 Participants agreed strongly that any researcher or official with access to the data should be appropriately vetted. A number of specific suggestions for possible vetting mechanisms were put forward, which included:

- a certified training course on data linkage and data security, supplemented with regular 'refresher' courses
- a comprehensive procedure, along the lines of a Disclosure Scotland check, for assessing whether an individual should be given access to data
- an accreditation scheme which signifies that an individual and/or organisation has met all of the criteria required to enable access to the data
- the introduction of a duty on public bodies to ensure that all researchers and officials with access to linked data are fully vetted, trained and monitored.

4.36 Participants went on to suggest that the oversight body should be responsible for checking that applicants meet the required criteria and should be held accountable if access is granted to anyone who does not. As already mentioned, transparency – in the sense of the public being able to access information on anyone requesting access to linked data – was also considered important.

4.37 There was agreement among participants that it was important for strict access procedures to be formalised and that these should be explicit in the Principles for the avoidance of doubt. One participant suggested that these should also specify the grounds on which access would be denied.

Data sharing agreements and sanctions

4.38 Participants agreed that data sharing agreements should be implemented and strongly supported the imposition of sanctions if agreements are not followed. The consensus was that such agreements should explicitly set out the roles and responsibilities of each individual and organisation involved in data linkage in relation to each of the Principles, particularly those concerning data security. It was felt that this would create greater accountability among individuals and organisations.

4.39 Participants also suggested that data sharing agreements should contain contingency plans to cover data breaches or losses. Again, they felt these should be explicit so that each individual involved would know exactly what to do in the event of an issue arising. Some people went on to say that this would help to mitigate the impact on the individuals whose data has been compromised and minimise adverse media coverage.

4.40 Participants were very keen to see incorporated into the Guiding Principles a range of specific sanctions; firstly, to deter individuals or organisations who may consider stealing or selling data; and, secondly, to

ensure that any individual or organisation that does breach the Principles is punished appropriately.

> Somebody is not going to break security if they know the consequence…
>
> (Male, aged 35-49, Inverness)

4.41　There was a consensus that sanctions should be very strict to deter anyone who may consider breaching the Principles. Specific suggestions included:

- banning culprit individuals and organisations from using data linkage or accessing data in the future, including closing loopholes that may allow them to access data through different or new organisations
- ensuring individual culprits lose their jobs – this was considered a particularly strong deterrent for professionals whose careers depend on their being able to access data
- issuing fines appropriate to the scale of the breach and also the circumstances of the individual or organisation being punished
- issuing judicial sentences to individuals, including community service or prison, again depending on the scale of the breach.

4.42　As is evident in these suggestions, there was broad agreement that sanctions should apply to both individuals and organisations. It was felt that this would incentivise organisations to ensure that relevant employees are fully trained and monitored in respect of data linkage. However, participants were keen to point out that sanctions should be considered on a case by case basis to ensure that the correct culprits are punished; for example, so that an organisation is not punished for the actions of an individual employee if it can prove that it has taken every possible step to prevent breaches occurring.

Outstanding issues

4.43　The various concerns and issues raised by participants in respect of the Guiding Principles, detailed above, were reflected in their responses to the end of event questionnaires. Participants were asked which issues they felt it was most important to consider or resolve as the Framework is developed. The most common responses were ensuring data security (mentioned by 45% of those who completed the question), ensuring transparency/ keeping the public informed (31%) and improving trust (22%).

'Beyond 2011'

4.44 *Beyond 2011* is a project is being run by the National Records of Scotland to assess alternative options for producing population and socio-demographic statistics, including the use of administrative data sources.

4.45 The project was briefly outlined in the presentation as an area of work that might benefit from improved data linkage, and participants were asked for their views on this. There was broad support for the objectives of *Beyond* and it did not raise any new privacy issues; indeed, there was a sense in which several participants felt that the census was more intrusive than research given that individuals are chased up and sometimes fined for non-completion of their forms.

4.46 Ten years was generally considered too long a gap between censuses, with younger participants, in particular, pointing out that individuals' circumstances, and society generally, change much more rapidly than this meaning data soon becomes obsolete. Additionally, several participants recognised that the census is expensive to deliver and questioned whether it is worth the money given the identified shortcomings of the data.

> I filled in my census; two years ago was it? My situation's changed since then, so it doesn't give a true reflection.
>
> (Male, aged 18-34, Glasgow)

> I think the world moves too fast now. Ten years is far too long, far too long.
>
> (Male, aged 35-49, Inverness)

Further public engagement

4.47 There was a strong appetite among participants for ongoing public involvement in the development of the Data Linkage Framework. This was underpinned by two main considerations: Firstly, a widely held view that the Government generally needs to do more to consult the public on important issues in order to engender a greater sense of confidence in its decisions and policies; and, secondly, a perception that the events provided important reassurances about data linkage and related issues (particularly around data security), and that more public engagement could similarly allay potential concerns among the wider public, and also serve to counter any media scare stories that may emerge as the Framework is implemented.

4.48 Participants acknowledged that it could be difficult to recreate the events on a wider scale – and, indeed, that many members of the public may not be interested in engaging so proactively on the subject of data linkage – so suggested a range of alternative initiatives that the Government could pursue to at least keep the public informed about the Framework. These included: media advertising campaigns; the distribution of leaflets (both to

households and public places such as council offices); and the setting up of a dedicated website that could serve as a 'one stop shop' for everything members of the public might want to know about data linkage – including research studies that are underway – and that could be referenced in any advertisements or leaflets produced. Some of the older participants were keen to emphasise that the Government should use a variety of methods to communicate with the public so that people who, for example, don't have internet access, are not excluded from engaging fully with the issues.

> A paper pamphlet giving you the basic outline […] and then linking to, say, a website with more detailed information on it for anybody who really wants to know [more].

> (Male, aged 35-49, Glasgow)

4.49   Data from the questionnaires completed by participants at the close of the events reinforces the potential for information provision to impact positively on public perceptions around data linkage. Just over half of participants (52%) who responded to Q6 of the questionnaire (n=52) reported that the events had changed their views on data linkage or the use of data in research, with most specifying that it had increased their understanding of data linkage in general, or of the beneficial ways in which linked data can be used. Around one in five stated explicitly that the event had increased their support for data linkage and/or the holding of data and information by public bodies.

4.50   Still, there was also a common suspicion that consultations and other forms of public engagement are just "tick-box" exercises to "rubber stamp" the Government's plans and that the Framework will be implemented regardless of public opinion.

5   CONCLUSIONS

5.1   This research was commissioned as a means of gaining insights into the public acceptability of cross-sectoral data linkage that may inform the development of the Scottish Government's proposed Data Linkage Framework. Deliberative methods were chosen in recognition of the complexity and unfamiliarity of the topic and thus the need for participants to be appropriately informed in order to meaningfully consider the relevant issues.

5.2   Consistent with a recent commissioned report on public attitudes to the sharing of health data for research (Aitken 2011b), one of the clearest findings to emerge from the events was the readiness and ability of participants to engage with the subject matter – they spoke at length, and with some sophistication, about key issues of relevance to the Framework, providing a rich and actionable evidence base for future decision making.

Concerns and sensitivities about data linkage

5.3   There were a number of factors underpinning attitudes towards data linkage but chief among these were: general concerns about an encroaching "big brother society"; negative experiences of data collection and use by commercial actors; and scepticism around the trustworthiness of public bodies to look after data and use it appropriately. The latter was partly fuelled by high profile cases of data losses and a view that public bodies are active in selling data to commercial organisations. Directly addressing these various issues as part of any communications surrounding the ongoing development of the Framework may be an important first step in building public acceptability.

5.4   Thus, although participants broadly supported the overarching objectives of the Framework, they did have specific concerns on which they sought reassurance, which tended to centre around the questions of:

- who would oversee the operation of the Framework
- who would have access to linked data, and specifically whether this would include commercial companies
- how individuals' privacy would be protected; in particular to prevent unsolicited contact from commercial companies
- how the data would be kept secure
- where overall accountability would lie if linked data was lost or stolen

5.5   Additionally, there was concern that linkage could lead to increased negative "labelling" of individuals due to the potential for labels to carry across sectoral boundaries and result in individuals or groups experiencing discriminatory treatment or stigma in multiple spheres. Such concerns often led participants to express a view that public bodies should only be able to access data that is directly relevant to their work.

Specific sector to sector linkages

5.6   The above concerns were discussed in relation to data linkage generally. There was little explicit differentiation between particular sector-to-sector linkages; notwithstanding some sensitivity around linkage involving criminal justice data and the potential for this to result in individuals experiencing discrimination across multiple spheres.

Reaction to the draft Guiding Principles

5.7   Perceptions of the draft Guiding Principles were somewhat mixed. On one hand, there was a view that the Principles go some way to addressing the main areas of concern and provide reassurance that data linkage will be carried out appropriately and securely. On the other hand, the Framework in general, and the Principles specifically, were commonly considered to be too "vague" and therefore open to interpretation and manipulation by vested interests. In particular, it was felt that there was a lack of detail surrounding who would be on the oversight body and the sanctions that would be imposed for breaching the Principles.

Safeguards to maximise public confidence

5.8   Further discussion of the Guiding Principles led to the identification of a number of safeguards that could be implemented in order to maximise public confidence in the Framework. These included:

- a requirement that anyone applying to use linked data must provide a strong justification to a commission or panel as to why their research is in the public interest
- publishing all processes and procedures surrounding data linkage, as well as details of who is undertaking research using linked data, to enable the public to monitor activity taking place under the Framework
- establishing an oversight body - comprising highly qualified professionals and, potentially, lay members - with responsibility for granting or refusing data linkage requests, ensuring that the Principles are upheld, and administering sanctions as required
- establishing accountability by placing data linkage under ministerial remit or the auspices of an independent professional or senior civil servant
- ensuring that explicit consent is obtained for uses of data containing names or other direct identifiers, and that, in the process, clear parameters are set around what is being consented to (i.e. the type of research), with any divergence requiring further consent
- requiring that consent for the uses of data containing names or other identifiers be obtained from data subjects themselves or from their next of kin, and preventing any oversight body from granting proxy consent
- ensuring all electronic systems used by individuals and organisations with access to linked data meet a minimum security requirement that is reviewed and updated frequently

- ensuring all researchers and officials with access to linked data are appropriately vetted through mechanisms such as: a certified training course; an accreditation scheme; or an assessment scheme similar to Disclosure Scotland.
- imposing strict sanctions on individuals and/or organisations responsible for any breaches of the Principles (which might include bans on future access to link data, and the issuing of fines or judicial sentences) and specifying the range of possible sanctions within the Principles.

Support for the objectives of *'Beyond 2011'*

5.9　There was broad support for the objectives of Beyond 2011, which did not raise any new privacy issues. Generally 10 years was considered too long a gap between censuses and several participants questioned whether the census represents value for money given that the data soon becomes obsolete.

Attitudes to ongoing public involvement

5.10　The notable readiness and ability of participants to engage with the subject of data linkage provides in and of itself a strong case for ongoing public engagement in the development, implementation and monitoring of the Data Linkage Framework. Across the events, there was a strong appetite for such engagement, which was underpinned partly by a perception that Government generally needs to do more to consult the public on important issues, and partly by a view that the events provided assurances about data linkage and related issues. There was specific support for a media advertising campaigns; the distribution of informational leaflets; and the setting up of a dedicated website that could serve as a 'one stop shop' for everything members of the public might want to know about data linkage. Older participants emphasised the importance of adopting a *multi-strand* public engagement strategy to maximise its reach.

Further research

5.11　While the research findings provide a clear indication of public sentiment around the issue of data linkage, they also point towards areas for further investigation. Firstly, it would seem worth attempting to assess the extent to which the various issues and concerns raised by participants are reflected among the wider population and to explore the relative perceived importance of the suggested safeguards. This would best be done through quantitative methods and, in particular, techniques that allow for the exploration of multiple considerations and trade-offs, such as discrete choice experiments.

5.12　Other issues that it would be useful to investigate further, whether qualitatively or quantitatively, are conceptions of the public interest; preferences around the composition of oversight bodies and implications of this for public trust; and the feasibility and acceptability of different approaches to obtaining consent for record linkage.

BIBLIOGRAPHY

Aitken, M (2011a) *Linking Social Care, Housing and Health Data: Social Care Clients' and Patients' Views*, Edinburgh: Scottish Government Social Research [online at: http://www.scotland.gov.uk/Publications/2011/09/20085846/0].

Aitken, M (2011b) *SHIP Public Engagement: Summary of Focus Group Findings*, Scottish Health Informatics Programme [online at: http://www.scot-ship.ac.uk/sites/default/files/Reports/Focus_Group_Findings_Briefing_Paper.pdf].

Bhopal, R., Fischbacher, C., Povey, C., Chalmers, J., Mueller, G., Steiner, M., Brown, H., Brewster, D.H. and Bansal, N. (2011), 'Cohort Profile: Scottish Health and Ethnicity Linkage Study of 4.65 million people exploring ethnic variations in disease in Scotland'. *International Journal of Epidemiology,* Vol.40, No.5, pp.1168-1175

Damschroder, L.J., Pritts, J.L., Neblo, M.A., Kalarickal, R.J., Creswell, J.W. & Hayward, R.A. (2007) 'Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records' *Social Science and Medicine* 64: 223-235.

Edinburgh Study of Youth Transitions and Crime: http://www.law.ed.ac.uk/cls/esytc/

Haddow, G., Laurie, G., Cunningham-Burley, S. & Hunter, K.G. (2007) 'Tackling Community Concerns about Commercialisation and Genetic Research: A modest interdisciplinary proposal', *Social Science and Medicine* 64, 272 – 282.

Heath, J. (2010) 'Emerging Consumers View of Secondary Uses of Medical Data', *IEEE International Symposium on Technology and Society*

Hunter, I., Whiddett, R.J., Norris, A.C., McDonald, B.W., and Waldon, J.A.(2009), 'New Zealanders' attitudes towards access to their electronic health records: Preliminary results from a national study using vignettes', *Health informatics Journal,* Vol 15, No.3, pp.212-228.

Jones, P. & Elias, P. (2006) *Administrative data as a research resource: a selected audit*, ESRC National Centre for Research Methods [online at: http://eprints.ncrm.ac.uk/452/]

Morris, A.D., Boyle, D., McAlpine, R., Emslie-Smith, A., Jung, R.T., and MacDonald, T. (1997), 'The diabetes audit and research in Tayside Scotland (darts) study: electronic record linkage to create a diabetes register'. *British Medical Journal*, 315:524

Scottish Government (2012) *A Scotland-wide Data Linkage Framework for Statistics and Research: Consultation Paper on the Aims and Guiding Principles*, Edinburgh: Scottish Government.

Trinidad, S.B., Fullerton, S. M., Bares, J. M., Jarvik, G.P., Larson, E.B. & Burke, W., (2010) 'Genomic research and wide data sharing: Views of prospective participants' *Genetics in Medicine* 12(8): 486-495.

Willison, D.J., Schwartz, L., Abelson, J., Charles, C., Swinton, M. Northcup, D. & Thabane, L. (2007) 'Alternatives to Project-specific Consent for Access to Personal Information for Health Research: What is the Opinion of the Canadian Public?' *Journal of the American Medical Informatics Association* 14(6): 706-712.

APPENDIX A: PROFILE OF WORKSHOP PARTICIPANTS

Criteria	Location			
	Stirling	**Inverness**	**Glasgow**	**Total**
	n	*n*	*n*	*n*
Total attended	24	22	27	73
Gender				
Male	12	10	13	35
Female	12	12	14	38
Age				
18-34	6	8	11	25
35-49	6	7	10	23
50+	12	7	6	25
Working status				
Working full or part time	16	14	20	50
Not working	8	8	7	23
Social grade classification				
ABC1	12	10	15	37
C2DE	12	12	12	36
Ethnic minority group				
Yes	2	2	3	7
No	22	20	24	66
Disability				
Yes	8	3	3	14
No	16	19	24	59
Children under 16				
Yes	9	6	4	19
No	15	16	23	54

APPENDIX B: WORKSHOP MATERIALS

Topic guide

Deliberative citizen engagement event on
'*public acceptability of cross-sectoral data linkage'*

9.30am – 10am: Arrival
- Registration, provide participants with name badge indicating which of the small groups they will be in
- Workshop facilitators and presenters to mingle
- Poster boards with agenda and purpose of the event
- Teas and coffees

10:00-10:15: Initial Plenary
- Thank respondents for attending
- Introduce Ipsos MORI and University of Edinburgh and the people involved and their roles during the event
- Explain, why they were invited, briefly explain the background to the study and emphasise that the focus is on the use of data, gathered by different public bodies and through particular studies, for research and statistical purposes only. Briefly outline agenda for the day.
- Explain key terms: data/information; linkage; research and statistical purposes. Clearly separate individual level data relating to service use/need from research and statistical purposes.
- Emphasise: confidentiality, no right or wrong answers, give everyone opportunity to speak, important to hear public views and involve public in these issues
- Permission to record discussions
- Housekeeping: toilets, fire exits, refreshments, mobile phones switched off

10.20-11:15: Initial warm-up session (55 minutes)
Aims: to explore, without prompting:
-General attitudes towards data collection and use
-Perceptions of how the SG and other public bodies might use data (particularly issues of trust)
-Initial reaction to concept of data linkage
-Concerns about particular data sources, such as medical records
[This is a warm up session, so needs to be quite open ended, yet not roam too far away from our central purpose. We will have clarified at the outset what we mean by data/information; data linkage; research and statistical purposes, so that this can set the overall boundary for the discussion]
- How do you feel about public bodies [USE CARDS TO IDENTIFY RANGE OF PUBLIC BODIES] collecting and holding information about you? [RECORD ANSWERS ON FLIP CHART]
[NB SG holds very little data about individuals as generally not delivering services. UK Govt eg HMRC and DWP do as do LAs.]
- Do you feel differently about different organisations?
  - To what extent do you trust these various bodies to look after such data? Probe issues of trust e.g. why do you say that
  [RECORD ON FLIP CHART; CREATE HIERARCHY IF POSSIBLE]
- How do you think such information is used?
  Probe: Do you feel differently about different uses? [RECORD ON FLIP CHART; IDENTIFY ACCEPTABLE/UNACCEPTABLE USES]
- How do you feel about how information about you might be used for research and for producing statistics about the population or parts of the population?

- What are your thoughts about who might want to do such research and why?
  - Probe types of research and researchers; see if public interest arguments emerge

EXPLAIN: As you have just heard in the introduction, the SG is interested in facilitating the linking of data that they and other public bodies hold about individuals in order to use existing information more efficiently (and therefore save money) and increase the range of analysis that can be done (compared to single datasets). This will help research which in turn should help provide societal benefit (such as better policies and services). It will also help produce better statistics on the Scottish population.

- What do you think about the idea of using information held by public bodies(for example, relating to education, health, housing) and research studies (such as social and health surveys) that have been linked in order to do further research or produce more information on the Scottish population?

  - Probe views on benefits/disbenefits

[USE STICKY WALL– this will enable people to write down their views initially – facilitator can then 'analyse' – identify themes]

- What do you think this type of linked information could be used for?

  - Probe what it should be used for

[USE STICKY WALL]
- Are there some types of information that you feel differently about (probe about different sectors – education, health etc) [USE CARDS TO IDENTIFY RANGE OF AREAS]

- What about data from commercial companies (for example, energy data; club card data?)

- How do you think SG might use this information and how do you feel about that? What about other public bodies, such as local authorities, police, universities? [USE CARDS AGAIN TO IDENTIFY RANGE OF SECTORS]

- How do you feel about commercial access to linked information held by public bodies?

End session with participants listing issues they'd like further clarification on (we can then attend to that in the presentation and Q and A).

11:15-12:00 Presentation (20 min) followed by Q&A (25 min)
At the end of the presentation, participant will be given a summary of the guiding principles to refer to over lunch
Q and A will take place initially with individuals working in twos/threes, highlighting questions and then in plenary. Facilitators will add questions that have arisen in their groups that have not been asked in plenary.
12.00-12.40 LUNCH
Poster Boards will have easy read versions of the framework in large print. Participants will also have individual leaflet in plain English. A poster with the presentation on will also be shown; this will include case studies of useful research that has used linkage.

12:40-13:45: Post- presentation breakout groups (65 mins)
Aims: to discuss participants' perceptions of the proposed strategy, including their:
- Overall concerns and more specific concerns regarding particular sector-to- sector linkages
- The extent to which the 'Guiding Principles' provide appropriate governance frameworks
- Developing recommendations for safeguards to address concerns and maximise public confidence
- Exploring support and concerns surrounding the Beyond 2011 project
EXPLAIN: You have had the chance to hear about the Scotland wide data linkage framework and also to ask questions and raise concerns. Do you have any further questions of clarification before we go on? I'd now like to ask you now to discuss the proposed strategy a little more

- Do you have any general concerns about it (consent, confidentiality, uses/abuses?)

- Do you have specific concerns (for example about linking particular types of information or linkages, uses/abuses)?Do you feel confident about the safeguards that have been put in place to ensure privacy (data controllers, PAC/S)?

- You have also heard something about the Beyond 2011 project, where SG is exploring the use of alternative sources of data for Census type outputs. We'd like to ask you a little more about that. What do you think about using existing, linked information held about you in this way?

- How do you feel about using commercial data in this way (Such as data from energy companies or club card data?)

EXPLAIN: There are several guiding principles set out in the framework (reiterate these and have them on cards). Can we discuss these a little more before moving on as a group to develop your own recommendations?

- Do you think these principles deal adequately with public concerns?

- Do you think some of them are more important than others?

[SORTING EXERCISE WITH THE CARDS – THIS WILL PROMPT DISCUSSION ABOUT EACH KEY PRINCIPLE AND GET A SENSE OF PRIORITIES – SPARE CARDS WILL ENABLE PARTICIPANTS TO WRITE ADDITIONAL PRINCIPLES]

- If you were asked to advise SG on how best to deal with public concerns and safeguard information, what would be the most important things to you?

[POST-ITS/FLIP CHARTS TO LIST AND THEN RANK THESE]

EXPLAIN: Just to finish up this discussion – you have all had the chance to have your say today. More generally

- What do you think are the main things that SG can take out of today?
- Would you like more information about how information about you is used?
- Would you like to be kept informed about what happens next? If so, in what ways?
- Do you think the public should be more involved in this issue? If so, in what ways?

[FLIP CHART]

- Summarise key messages from the day and ensure agreements and disagreements noted
- Thank participants for their input and reiterate next steps (our report; and SG will probably consult further on the details of what they propose to do)
- Distribute post-event questionnaire and incentives

Informational presentation

Data linkage in Scotland
What it is and what is proposed

Dr Claudia Pagliari
and
Professor Sarah Cunningham-Burley
The University of Edinburgh

What is this event about?

- The Scottish Government has drafted a 'Framework' that sets out 'principles' for joining up (linking) information (data) collected by different agencies, such as the NHS, Local Authorities, Government.

- They want to know what you think about this and what, if any, concerns you might have.

- Understanding what the public thinks will help to inform these plans.

What it is NOT about
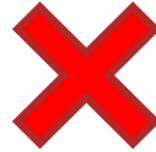
- Sharing of information about **named** individuals between agencies to enable joined-up <u>services</u>
  - e.g. to assess whether a child is at risk, whether someone is falsely claiming benefits, or to coordinate diabetes care between different health service providers

What it <u>IS</u> about

- The use of linked information to compile official *statistics*

This means information about the Scottish population or groups within the population, such as that produced by the National Records of Scotland. Linked information will give a more detailed picture of the Scottish population and factors that influence their lives.

Official statistics are made public so anyone can look at them, but these are always in the form of 'aggregate' information with nothing that can reveal an individual's identity.

What it IS about

- The use of linked information and data for *research*

This means, for example, using linked information to create a large study, such the Scottish Longitudinal Study. This uses data from Scottish administrative and statistical sources. The data are anonymous – no names and addresses are attached to it.

It can also mean linking information held by different agencies to surveys, such as the Scottish Health Survey or the Growing Up in Scotland study. Those participating in these surveys give consent for this to happen. Sometimes researchers may to link data in order to answer an important research question, such as in the Scottish Health and Ethnicity Linkage Study

Remember this guy?

Balancing public interest and public good with individual privacy and individual needs

What are your views? Are you concerned about your privacy?

"The needs of the many outweigh…
The needs of the few…or the one"

Mr Spock in Star Trek II, The Wrath of Khan, 1982

How is IT done?

- Data linkage is about mapping information about the same individuals across different systems and agencies

➢ This information is mostly stored as numbers (e.g. visits to hospital, benefits claimed, your age, exam grades)

➢ Before linkage is done, names and addresses are removed and replaced with anonymous codes consisting of numbers and letters

➢ The linkage is done by matching pieces of information that are labelled with the same code in the various datasets

So what's new here?

- More computerised data are becoming available and the new knowledge we are able to gain is likely to be immense
- To make best use of these data for public benefit, we need better ways to organise it, link it and safeguard privacy
- The new Framework outlines principles and methods of ensuring safe use and it also proposes a national coordinating centre for data linkage
- This aims to promote public benefit whilst protecting personal privacy

What do you see as public benefit? Does it challenge personal privacy or personal benefit?

What are the benefits?

- Better information will help develop the best services to meet our needs

  For example, we now know that Looked After Children are 8 times more likely to be excluded from school than other children. This is because information from many sources were linked together. Support can now be targeted at these children

What are the benefits?

- Produce better statistics

  Linking administrative data can help provide more up to date information on the population than the ten yearly Census does.

  Linked data can also give a more detailed picture (for example sharing social work and education data)

What do you think about this amount of information being held and used?

What are the benefits?

- Lower cost research tracking the population over time

The Scottish Longitudinal Study has shown that there is an increased risk of mortality (death) due to widowhood. This means that it is important to target health care to this group

What are the benefits?

- Better ways to evaluate new health, social, educational or other programmes

  For example, clinical trials can be extended by included linked data once the trial is over.

  Large scale biobanks, like Generation Scotland, have asked for consent to link further data on the participants. This will help understand important issues such as healthy ageing.

Would you give your consent to such a thing?

What are the benefits?

- Scotland has a world wide reputation for its data linkage work in health research

  Building on this success, with cross-sectoral data linkage, will make Scotland an excellent place to do research.

Do you think that is for public benefit?

Why do we need a new framework?

It is argued that it is in the public interest to safeguard individuals' right to privacy and to make efficient use of data for statistical and research purposes.

- The Data Linkage Framework aims to help everyone involved in data linkage to balance these values
- To provide the public with assurance that personal data are being used safely and appropriately
- To provide a common framework of principles and standards
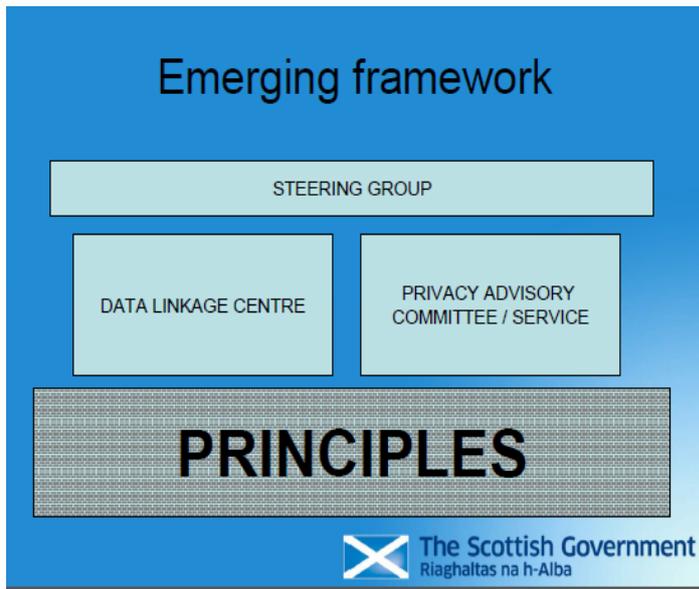
What are the principles?

- Public Interest
- Governance and Public Transparency
- Privacy
- Removal of names and direct identifiers
- Consent
- Security
- Access and Personnel
- Data sharing agreements and sanctions

Time for questions and discussion

- What are your first thoughts about these issues?

- Do you think such cross-sectoral data linkage is a good idea?

- Do you have any concerns about these issues?

- Do you trust the people and agencies involved?

Summary of Guiding Principles used to facilitate group discussions

Guiding Principles

Public Interest
This refers to the benefits that data linkage may bring through better research and statistics and the need to balance this against individual rights and privacy protection

Governance and Transparency
This refers to the importance of being open about the ways in which data linkage is happening and about the safeguards that are in place

Privacy
This refers to a strong commitment to privacy protection, for example by ensuring individuals cannot be identified at any stage

Removal of names and direct identifiers
This refers to a commitment to using data without names and other information that identify someone; any departure from that must be justified and approved

Consent
If personal data where an individual may be identified is to be used, then that person must give consent. If that is not possible or practical, then permission is required from an oversight body

Security
All data linkage must conform to the highest standards of security and follow appropriate policies

Access and Personnel
An access policy must be developed and all those using linked data must be approved and keep to a Data Sharing Agreement

Data sharing agreements and sanctions
All those involved in data linkage must stick to a clear agreement. If they do not do so, then sanctions will come into force

Post-workshop questionnaire

Data Linkage Strategy Discussion Event
Feedback Form

We are interested in your opinion of today's event. Please take a few minutes to complete this short questionnaire.

Q1 How far do you agree or disagree with the following statements?
PLEASE TICK ✓ ONE BOX ONLY FOR EACH ROW

		Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
a)	I found the event enjoyable	☐	☐	☐	☐	☐
b)	The presentation was informative and interesting	☐	☐	☐	☐	☐
c)	The opportunity to ask questions was useful	☐	☐	☐	☐	☐
d)	Overall I feel better informed about the material discussed..................................	☐	☐	☐	☐	☐
e)	The design of the event was stimulating..	☐	☐	☐	☐	☐
f)	There was enough time to hear the views of others	☐	☐	☐	☐	☐
g)	There was enough time to share my views with others....................................	☐	☐	☐	☐	☐

Q2 What, if anything, did you enjoy most about the event?
PLEASE WRITE IN BELOW

Q3 What, if anything, did you enjoy least about the event?
PLEASE WRITE IN BELOW

Q4 What, if anything, would have made the event better?
PLEASE WRITE IN BELOW

```




                                48


```

Q5 Is there anything you would have liked to have said but didn't?
PLEASE WRITE IN BELOW

```



```

Q6 Did this event change your views about any aspect of data linkage and/or the use of data in research?
PLEASE TICK ✓ ONE BOX ONLY

Yes   No
☐   ☐

If **yes**, please tell us how your views have changed? **PLEASE WRITE IN BELOW**

```



```

Q7 Thinking about everything that has been discussed today, what do you think are the most important issues to consider or resolve when developing the data linkage framework?
PLEASE WRITE IN BELOW

```
 
 
 
 
 
 
```

Q8 Did you find today's venue suitable or not suitable?
PLEASE TICK ✓ ONE BOX ONLY

Suitable	Not suitable	Don't know
☐	☐	☐

If you found the venue unsuitable, why do you think it was not suitable? **PLEASE WRITE IN BELOW**

```
 
 
 
 
 
```

Thank you very much for the feedback.
Please hand your completed form to one of the moderators.

GOVERNMENT SOCIAL RESEARCH

Social Science in Government