SHIP Privacy Impact Assessment
A document of core programme 1

Background

Scotland has very comprehensive health service data of high quality and consistency. However it is still rather disparate. A great deal of meaningful research can be done when data from different sources is brought together. Fortunately the almost universal use of the Community Health Index (CHI) number means that Scotland's health data can now be linked to allow patient based analysis and follow up. By adding the CHI number to non-health datasets ("CHI seeding") there is also the capacity to conduct health / non-health linkages to gain insights into health and social care or health and education for example.

Research using linked health data is currently possible and has yielded many beneficial findings but is difficult to achieve. This is largely due to the complex legal environment. The lack of clarity has led to a conservative approach by many data controllers. In addition the convoluted and unco-ordinated system of approvals has meant frustration and delay for researchers.

SHIP aims to make a system that is anonymous and secure enough to give data controllers comfort to allow their data to be linked for research and to streamline the approvals process for researchers.

SHIP will mostly deal with non-consented data that is gathered routinely by health care providers. Linkage will be done on a per project basis – i.e. data will be brought together to answer a particular question rather than a permanent merging of datasets. Linkages can be done that connects data about people from different sources without divulging anything about any particular individual.

SHIP will take a risk based approach to governance. Under the SHIP system the usual mode of access will be via data safe havens, as recommended by Walport and Thomas in their Data Sharing Review of 2008[1]. They define "safe haven" as an environment for population-based research and statistical analysis in which the risk of identifying individuals is minimised. Researchers will be able to access this secure environment either remotely or in person as appropriate. Walport and Thomas also recommend that a system of approving or accrediting researchers who meet the relevant criteria to work within those safe havens is established. They envisage researchers working in safe havens being bound by a strict code, preventing disclosure of any personally identifying information, and providing criminal sanctions in case of breach of confidentiality.

SHIP proposes to establish a National Safe Haven within NHS National Services Scotland (NSS). It is intended that the national safe haven be used for nationwide research, or studies that require linkage of datasets from multiple regions or that have multiple data controllers. This is within the scope of current funding.

It is also proposed that local safe havens will be established at the Scottish Academic Health Science Collaboration (SAHSC) nodes in NHS Grampian, NHS Greater Glasgow & Clyde, NHS Lothian, and NHS Tayside. Other local safe havens may follow.

[1] http://www.justice.gov.uk/reviews/docs/data-sharing-review-report.pdf

By providing a secure environment for research and a streamlined approvals process, more datasets should be accessible more quickly. Thus the hope is that SHIP will become a powerful resource for health research.

SHIP is a collaboration of the Universities of Dundee, Edinburgh, Glasgow and St Andrews with the Information Services Division of NHS Scotland. In developing these plans SHIP has consulted widely. Representatives of the following organisations have attended SHIP events:

Affiliation
AstraZeneca R&D Science Policy UK
BHF Glasgow Cardiovascular Research Centre
Biomedical Research Institute, University of Dundee
Centre for Population Health Sciences, Edinburgh
Centre of Academic Primary Care, Aberdeen
Chief Scientist Office
City University London
Clinical & Population Sciences and Education, Dundee
CSO Public Involvement Group
Department of Health
ESRC Secure Data Service, UK Data Archive
General Register Office for Scotland
Generation Scotland
GlaxoSmithKline
Health Informatics Centre, Dundee
Institute for Digital Healthcare, University of Warwick
Medical Research Council
NHS Central Register
NHS Greater Glasgow & Clyde
NHS Information Services Division
NHS National Services Scotland
NHS Tayside
Nursing Midwifery & Allied Health Professions Research Unit, Stirling
Ontario Agency for Health Protection and Promotion, Toronto, Canada
Robertson Centre for Biostatistics, Glasgow
Scottish Government
Scottish Primary Care Research Network, Dundee
Scottish School of Primary Care, Dundee
University of Aberdeen
University of Dundee
University of Edinburgh

University of Glasgow
University of Helsinki
University of St Andrews
University of Swansea
Virtual Microdata Laboratory, Office for National Statistics
Wales Office of Research and Development

In addition SHIP representatives have attended meetings of the Directors of Public Health, the Association of Research Ethics Committees, the CHI Advisory Group and the Caldicott Guardians Forum. Members of the SHIP Management Committee have direct links with the Privacy Advisory Committee and the Administrative Data Liaison Service.

Public engagement is one of the core programmes of SHIP, and the team has conducted a series of focus groups to explore citizens’ reactions to SHIP plans.

It is envisaged that SHIP can deliver a system that will provide a streamlined experience for users and more be cost effective than current systems.

Design of the Infrastructure

The Infrastructure

The SHIP Research Infrastructure is designed for the provision of linked datasets. It is anticipated that researchers will apply directly to data controllers for access to non-linked data, i.e. extracts of single source datasets. However data controllers may request that researchers use a safe haven facility to analyse their extract.

The detail of how the data is handled will depend on whether use falls within any consent associated with it and on the privacy impact assessment conducted as part of the proportionate governance process (see section 7). It should be noted that there may be occasions when data controllers will be happy to release innocuous data to the researcher directly, outwith a safe haven environment.

The infrastructure has three components which will work together to provide a timely and consistently high performance research service:

i. **A SHIP indexing service** will maintain a population index based on a unique patient identifier (UPI; eg the Community health Index (CHI) in Scotland). The indexing service will add anonymised identifiers (referenced to UPI) to individual records for the purposes of linking these records across two or more datasets. The indexing service will be separate from the linkage agent.

ii. **A SHIP linkage agent** will use anonymised identifiers to perform the matching of records belonging to individuals from two or more datasets to form a single linked dataset. The identifiers for the linkage will be provided by the indexing service.

iii. **SHIP safe havens.** These have three key characteristics, as defined by the Thomas/Walport Data Sharing Report:

- The safe haven will provide a secure environment for the linkage, storage and analysis of personal data.

- Access to data within safe havens may be from a dumb terminal within the safe haven or remotely via secure thin client technology dependent on risk assessment.

- Only ‘approved researchers’ will be permitted to access the data and they will be bound by a strict code, which prohibits disclosure of any personal identifying information. Safe havens will carry out statistical disclosure control on outputs to prevent accidental disclosure. The extent and level of disclosure control checks for a given project will be agreed with data controllers.

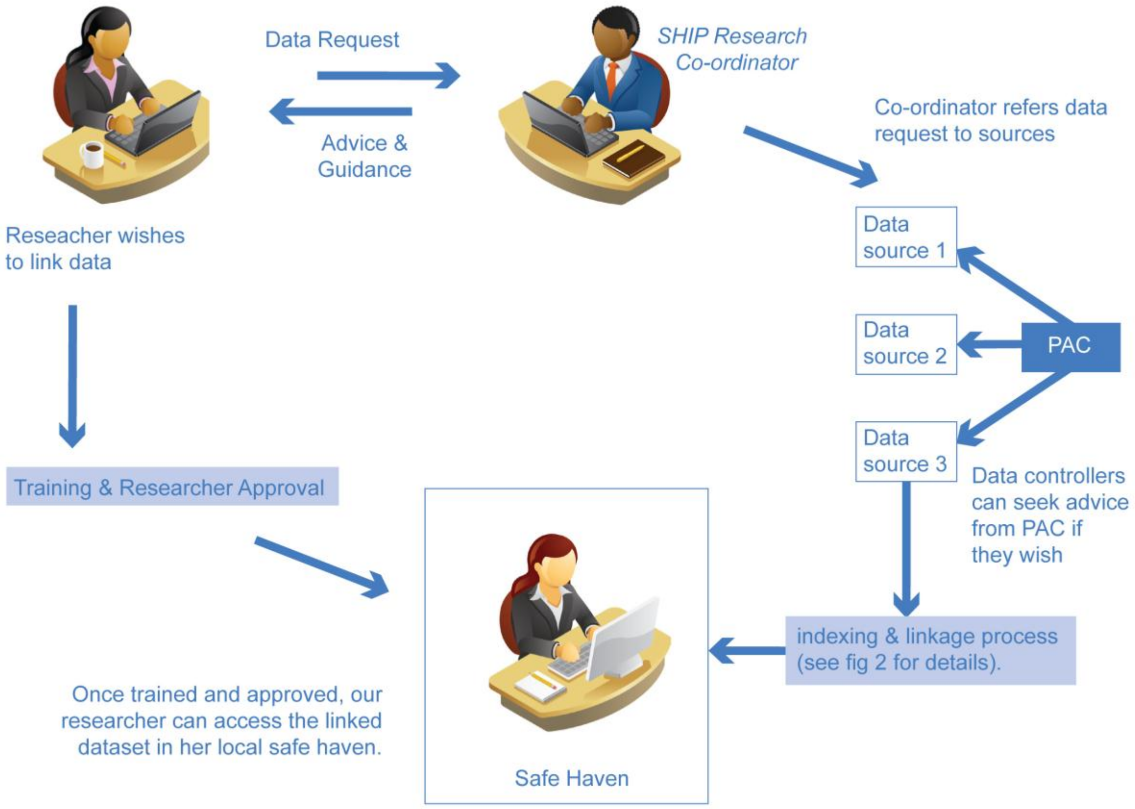
Data Request
SHIP Research Co-ordinator
Advice & Guidance
Co-ordinator refers data request to sources
Reseacher wishes to link data
Data source 1
Data source 2
PAC
Data source 3
Data controllers can seek advice from PAC if they wish
Training & Researcher Approval
indexing & linkage process (see fig 2 for details).
Once trained and approved, our researcher can access the linked dataset in her local safe haven.
Safe Haven

Figure 1: Overview of the SHIP infrastructure.

To ensure high levels of information security and the protection of subject confidentiality the storage of contributory datasets, indexing, linkage of data, and storage of the final dataset will be carried out separately. In practice this means that no individual should be directly involved in any more than one of these processes, but a single organisation could host more than one activity with appropriate segregation of roles and IT facilities. The Indexing Service will be 'stand alone', because this is the only function for which patient identifiers are required. The Safe Haven will be responsible for the remainder of the processes, which use anonymised data: linkage of data, provision of analytical software, the separate storage of the source and linked datasets and the analytical outputs. The Safe Haven will also be responsible for the implementation of other key functions, including an inventory of datasets with associated metadata, statistical disclosure control, accreditation of researchers (as part of a central register of approved researchers) and adherence to the good governance framework.

The Linkage Process

The indexing service receives only a project code, local identifiers and subject identifiers from the data sources for each of the datasets that are to be linked and no other data. The indexing service creates a study specific anonymised identifier for each subject (called a study number) and returns this with the associated project code, the local identifier and a score describing the likelihood that the linkage is correct. The study number for any individual subject will be different for each submitted dataset in which (s)he appears, to minimise the risk to subject confidentiality. The indexer will also supply a linkage key to the linker in the Safe Haven so that the datasets can subsequently be joined together by matching up the study numbers. The linker receives no other information from the indexer. This process is shown pictorially in Figure 2 below.

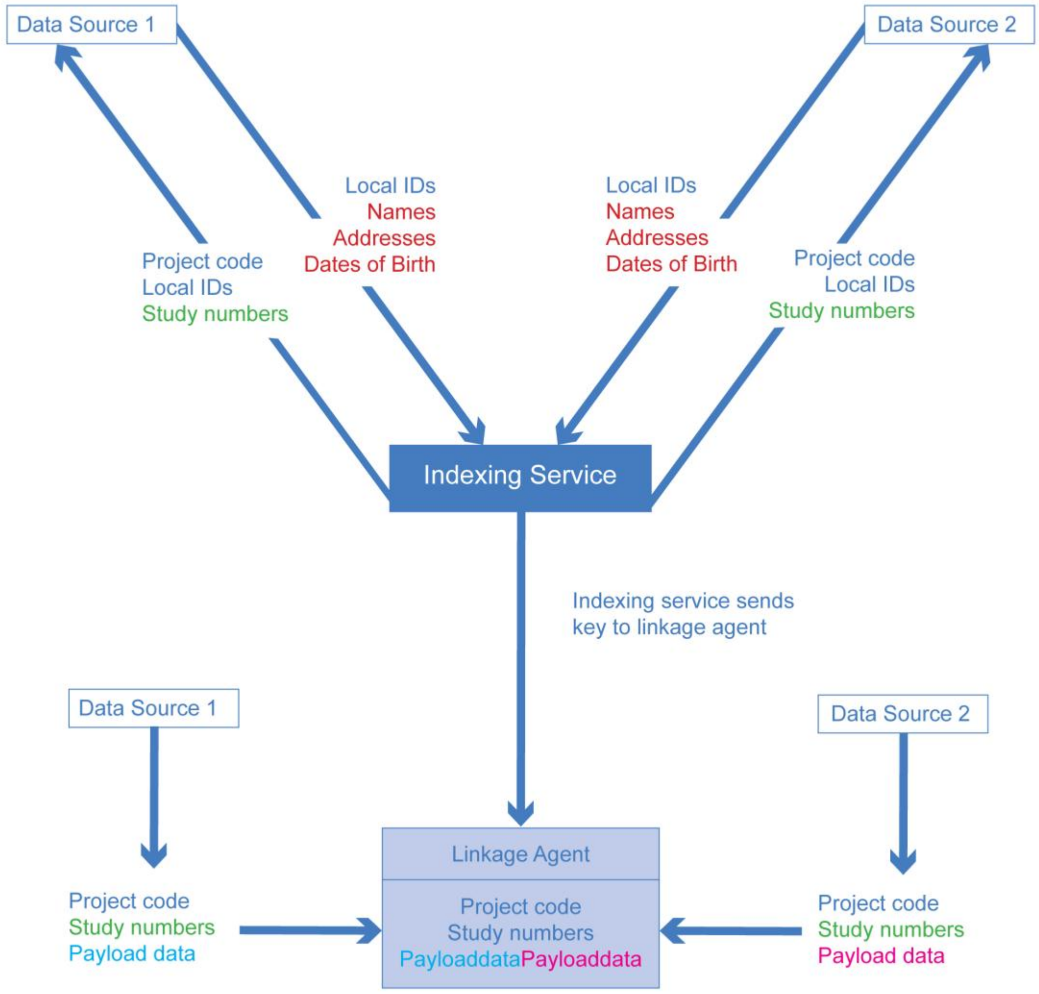

Figure 2: The linkage process: the indexing service uses demographics to assign project specific identifiers (study numbers) which are returned to the data sources in encrypted form. The key to decrypt the study numbers is sent to the linkage agent. Then the data sources send anonymised data to the linkage agent identified by encrypted study numbers only. The linkage agent decrypts and links the data on the study numbers.

The linkage agent and secure access facility will lie within NHS net. The indexing service must be able to receive and transmit information across the NHS net. Data may be supplied by other secure modes of transmission if these comply with SHIP governance principles. All information must be encrypted before transmission between data controllers, safe havens, indexing and linkage services. The Linkage Agent receives the study code and the study numbers together with the information the researcher needs ("payload data") from the data sources, as approved by the data controller(s). The Indexing Service supplies the linkage key so that study numbers can be matched across datasets, as described above. The Linkage Agent does not receive identifiable information (e.g. names, addresses or CHI numbers). The payload will generally consist of a subset of fields from any given dataset, being those that the researcher requires and for which permission has been obtained. The Linkage Agent uses the linkage key received from the Indexing Service to join datasets for the study and deposits the linked dataset in a separate area of the Safe Haven. The Safe Haven is the Data Controller for both the received datasets and the newly created linked dataset.

SHIP safe havens will follow the approach of the Office for National Statistics developed for their Virtual Microdata Laboratory. The safe haven holds the linked datasets and ensures that only approved researchers can gain access. Researchers will access the data held within the Safe Haven either remotely or via a dumb terminal in a secure access facility dependent on risk assessment. The secure access facility may be situated either within the same safe haven that holds the data or in another safe haven (with appropriate permissions). Analytical software will be available within the safe haven for use by researchers. The dumb terminals will be configured so that the researcher cannot download or remove any of the data or outputs held at the Safe Haven. A dedicated file space will be provided for the researcher to store their outputs pending release by the safe haven. De-identified data will be held separately from any data that carry identifiers eg consented datasets. Every Safe Haven must keep a suitable record of the use of its facilities for security and audit purposes. Software is available to provide a full log of terminal access.

The Safe Haven is responsible for undertaking Statistical Disclosure Control prior to release of analytical outputs to researchers. This will be done by appropriately trained employees of the safe haven. Once the output is deemed safe it will be sent to the researcher electronically. The level of disclosure control required will vary between studies. It is the responsibility of the data controllers for the contributory datasets and the Caldicott Guardians to decide upon the appropriate level of disclosure control at the beginning of the project before the datasets are linked and access is provided to the researcher.

The Safe Haven will provide an archiving service for all linked datasets so that researchers can return to the dataset for an agreed specified period of time following the initial analysis. While an extension to the time may be easily arranged, the analysis must still relate to the research question in the original application. If not then another application must be submitted. It is recognised that a well-developed research dataset is a valuable resource that may be useful over a long lifespan. Automatic destruction may not always be appropriate.

Analysis of the privacy issues arising

Key characteristics

Several aspects of the SHIP system mean that the privacy risks must be thought about carefully and taken very seriously. The data involved is sensitive personal data regarding health and it will be handled in new ways. New technologies are being developed for handling data in safe havens and there will be a much greater capacity for linkage.

There will be increased potential to bring together data on one individual from different spheres, possibly in large volume about each individual. This makes it doubly important that this is done either with the consent of the individual or anonymously with authorisation from an appropriate body. SHIP goes to great lengths to ensure anonymisation although it is recognised that the more details that are collated about a person the less anonymous they become.

There will be increased potential to bring together data about a large number of individuals. This has enormous benefits in the statistical power of the research that can be done. And provided the technical infrastructure is designed for the required capacity this does not in itself pose a particular risk to privacy – quite the reverse.

The system will use an existing identifier – the Community Health Index (CHI), as it is now used almost universally throughout the health service in Scotland. Any non-health datasets will be seeded with CHI for linkage. However the CHI number will not be given out to external data sources so they will not have the means to gain any new information about their clients. They will however have to provide lists of who their clients are to the indexing service. So another organisation will get to know who is "on their books". Since no information *about* those clients is being given to the indexing service this is not expected to pose a problem.

The potential to bring together data from multiple agencies needs particular focus. Data could be brought together from NHS, academia and private sector service providers opening up rich seams of research. However this could be seen as the breakdown of personal data and identity silos. SHIP addresses this concern by ensuring that demographic data (the identifiers) are always kept separate from the payload data (characteristics of the individual). The indexing service handles the demographics but not the characteristics and the linker/safe haven handle the characteristics but not the demographics. For linkages using the national safe haven this means that demographics and characteristics are handled by the same organisation, National Services Scotland (NSS), so external data controllers must have trust in NSS resolve to keep the two halves of their dataset apart. NSS must be able to give comfort to data controllers regarding the effectiveness of their "Chinese walls". Given the excellent record that NSS has in regard to data protection this should be easy to demonstrate.

Similarly it is vital that safe havens keep data with identifiers (eg consented data) separate from anonymous datasets. Otherwise the anonymous data could be re-identified using pattern matching. They also need to be able to demonstrate robust systems for ensuring this separation.

The linking of personal data from multiple sources does however give rise to issues of data quality, and semantic interoperability. So fields that appear to hold the same kind of data may not. This is because having been collected for different purposes, the data may have been collected in different ways and therefore mean something slightly different. This would have implications for the research and so an ontological mapping exercise may be needed.

The objective is to allow researchers to work on potentially disclosive data and the focus has been on the creation of systems and environments to allow this to happen with minimal risk (see section on Design of the Infrastructure). However in many cases researchers do not need and therefore will not receive identifying fields. Hence for much of the time there will be negligible risk from researchers. The much greater risk is from within NSS and the other centres that do indexing and linkage because the workers involved in these processes have the opportunity to access wide ranging sensitive multi-source data on identified individuals simply by contravening local standard operating procedures. Consequently the trustworthiness of staff is one of the keystones on which the structure rests. However these are not new issues and the organisations in question have systems in place to minimise these risks.

To reframe our analysis let us consider the various actors and what they can find out about data subjects:

Data sources can find out no new information about their clients. Although the CHI number is used for linkage data sources never see CHI numbers. It's important that the indexing service issues different project specific identifiers to each data source for the same person, otherwise data sources could gain new information on their clients from each other, by collaboration or stealth.

Indexers get to know who is known to which data sources but no detail about individuals.

Linkers, safe havens may see a great deal of information but without identifiers. That said they may get so much information that it becomes potentially disclosive and here standard operating procedures that support privacy and the trustworthiness of staff are relied upon.

Researchers may also see a great deal of information but without identifiers - so much information that it becomes potentially disclosive. The safe haven environment will mitigate this to some extent – i.e. the difficulty in removing any data. But there is always what the researcher can see and remember and here training, trustworthiness and sanctions come in to play.

Thus it is clear that privacy is preserved reasonably well provided the functions are distinct. Problems arise when indexers are also linkers or researchers are also data sources etc.

Business case

ISD currently completes approximately 50 projects per year. The new infrastructure is designed to cope with up to 300 projects. Under the current pricing structure the break even point for SHIP will be 130 projects (assuming average charge per project is the same as that charged in 2010/11). If charges are increased to £500 or £750 per day the break-even points are 90 and 60 projects respectively.

On this basis proposed SHIP model appears to be potentially more cost effective than the current MRL and Indexing Service, mainly because of the increased capacity

However, the decision regarding whether or not to proceed with the development of the SHIP Safe Haven model (Indexing and Linkage) should not be made purely on the ground of the relative costs of the current and future arrangements; although this is obviously an important factor. Other issues to be taken into account include:

a) NSS has a commitment to the funders of SHIP as laid out in the original grant application

b) The SHIP system will provide a much more secure way to handle linked data. Currently, linked datasets are given to researchers once anonymised. The SHIP model will hold data on a secure server, with remote access. Only ‘approved researchers’ will have access to data and statistical disclosure control will be carried out within the Safe Haven

c) The Indexing System will improve the speed and efficiency of linkages that ISD has to undertake as part of core business. The Indexing Service will have a greater capacity and efficiency than the current system.

The Indexing and Linkage Service has potential benefits beyond meeting researchers’ needs and SHIP commitments. The system could support both the pharmacovigilance initiative and the types of data use envisaged within the Information Strategy. For example, ISD will run as a SHIP pilot a project to link health and social care data to study care pathways for older people. This is of relevance to the Information Strategy and also supports the SG’s policy for convergence of health and social care. Similarly, the availability of linkable patient based prescribing data will greatly enhance our capacity for research in polypharmacy, drug safety and other aspects of pharmacovigilance.

Design features adopted to reduce privacy risk

A Safe System

To ensure high levels of information security and the protection of subject confidentiality the storage of contributory datasets, indexing, linkage of data, and storage of the final dataset will be carried out separately. Data sources can do much of the selection to avoid giving out identifying information.

The fact that identifiers will be kept separate from characteristics data and the indexing service will be "stand alone". Characteristics data will be associated only with anonymised identifiers.

Much of the security relies on the security of NHS net since the linkage agent and safe haven will sit within it. However the indexing service must be able to receive and transmit information across the NHS net boundary. All information must be encrypted before transmission between data controllers, safe havens, indexing and linkage services.

Controlling who can work on the data

Only accredited researchers will be given access to the facilities. The accreditation process will involve verification that the researcher

- Is associated with an approved institution
- Has completed a course in this case the SHIP online training module
- Has agreed to be bound by a strict code which precludes disclosing personal information. There will be sanctions for breach of this agreement.

Controlling which data researchers see

Projects only go ahead with the permission of the data controller who has the benefit of advice from a National Privacy Advisory Committee. Where appropriate the project will also be dependent on a favourable ethical opinion from an Ethics Committee.

Researchers will not see identifying fields such as name address and date of birth. Even if a researcher needs location and age information this can be provided without overt identifiers. For example by supplying age bands and area codes rather than full postcode.

Controlling what researchers can do with the data

Researchers access the data via safe haven terminals that have all data saving devices disabled: no memory sticks; no CDs; no internet capability, thus preventing the unauthorised removal of data. Thin client technology (eg Windows Terminal Services or Citrix) makes it feasible to expand the safe haven to the user's own computer wherever they are. This allows centrally managed secure terminal or terminal-emulation for authenticated users. While allowing remote access in this way means less control over the researchers physical environment, permission for this would be based on a risk assessment on advice from PAC.

Controlling what information is released into the public domain

The statistical results produced may be checked by officers of the safe haven to make sure they do not contain any disclosive results. This is called statistical disclosure control and involves processes such as:
Table redesign: sometimes called recoding, this involves combining categories of row and / or column variables to increase the number of respondents in cells
Cell suppression: the values in low frequency cells are hidden (primary suppression). However where there are totals, the missing value can be found by subtraction from the row and column totals. This necessitates "secondary" suppressions of other values to prevent disclosure by differencing.
Rounding: Rounding introduces a degree of ambiguity to the data so that it is impossible to tell which values have been rounded. The table should be annotated as having been rounded so that it is clear that published zeros may not be true zeros.

Within the safe haven

De-identified data will be held separately from any data that carry identifiers eg consented datasets. Every Safe Haven must keep a suitable record of the use of its facilities for security and audit purposes. Software is available to provide a full log of terminal access.

The Safe Haven will be responsible for adherence to the good governance framework throughout.

Analysis of the public acceptability of the scheme

SHIP’s public engagement team have run a series of focus groups on a range of relevant topics and highlighted the enthusiasm and competence of members of the public to engage on this subject.

There were a range of considerations influencing attitudes towards data sharing and/or linkage, for example: what is the purpose of data collection/sharing; what is data used for; who has access to the data and; how is it safeguarded against misuse.

To summarise conclusions there seems to be broad but not unconditional support. Generally speaking public support is dependent on the research being of clear benefit to patients or society at large. It is also dependent on individuals having some measure of control over use of their data. As expected confidentiality was an important consideration, however, participants suggested that they may be happy for identifiable, or potentially identifiable data to be used so long as they had control over this and trusted the individuals/organisations accessing the data.

The focus groups also highlighted a perceived need for greater openness and transparency about how data is collected and used.

Consideration of all these issues are being woven into the emerging plans for the SHIP infrastructure.

Conclusion

In conclusion the project has considerable potential to benefit public health. The project involves disclosure of personal data (including non-health data) to the NHS. However NHS already works under tight privacy regulation. It is anticipated that the systems and procedures proposed will minimise the disclosure risk to acceptable levels in order to realise the benefits.

Appendix: Relevant Legislation

The lawfulness of the use of PII in Scotland depends on compliance with the Data Protection Act 1998, the Human Rights Act 1998 and the common law of confidentiality.

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data
Key points to note here are:

- Privacy protection is important but the sharing of data is also a key objective of the Directive
- Processing and sharing of data must comply with 8 key data protection principles laid down in the Directive and which must be embodied in domestic laws.
- Member States have a degree of flexibility in how they regulate the processing of data. In particular, consent to processing it not an absolute legal requirement and processing can be justified in the 'substantial public interest' without the need for explicit consent.

Human Rights Act 1998 (HRA)
Article 8(1) provides that the UK must recognise and respect the individual's right to respect for private life. Seeking consent for data processing demonstrates respect for an individual’s private life as would anonymisation of the data to minimise any likely harms.

However Article 8(2) provides that the right to a private life is not absolute, and should be balanced with other interests, such as the protection of health or the protection of the rights and freedoms of others.

So Article 8(1) can be intruded upon where it is necessary and proportionate to do so i.e. where:
a) it addresses a 'pressing social need'
b) its operation is proportionate and
c) the reasons advanced for its existence are 'relevant and sufficient'.

Where it is necessary to process data without consent in order to carry out socially useful research penalties for an Article 8 infringement could potentially be avoided by arguing that the encroachment upon the right to respect for private life is proportionate and necessary, the balance of benefits and harms being carried out by a responsible authorising body. However this has never been tested in court.

The common law duty of confidentiality
The common law duty of confidentiality precludes disclosure of confidential patient information. However, legitimate justifications for disclosing such information exist, including where consent has been obtained and where disclosure is in the public interest or where there is a legal requirement e.g. public health legislative requirement to notify a particular disease.

Data Protection Act 1998 (DPA)

The Data Protection Act 1998 specifically relates to "personal data" which the act defines as, "data which relate to a living individual who can be identified—

(a) from those data, or

(b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller,"

So it follows that where data is anonymised it does not fall under the Data Protection Act.

According to the Act, data that falls within its provisions must be processed according to eight principles:

1 Personal data shall be processed **fairly and lawfully** and, in particular, shall not be processed unless—

(a) at least one of the conditions in Schedule 2[2] is met, and

(b) in the case of sensitive personal data[3], at least one of the conditions in Schedule 3[4] is also met.

[2] *To paraphrase Schedule 2 - Conditions relevant for purposes of the first principle: processing of any personal data:* either the subject has given consent to processing or processing is necessary in one of several defined ways.

[3] *In this Act "sensitive personal data" means personal data consisting of information as to—*

a) the racial or ethnic origin of the data subject,
b) his political opinions,
c) his religious beliefs or other beliefs of a similar nature,
d) whether he is a member of a trade union
e) his physical or mental health or condition,
f) his sexual life,
g) the commission or alleged commission by him of any offence, or
h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.

[4] *To paraphrase Schedule 3 - Conditions relevant for purposes of the first principle: processing of sensitive personal data:*

- Subject has given explicit consent to processing
- Processing is necessary by law or for state reasons
- The processing is necessary (a) to protect the vital interests of the subject or another person, where—
  (i) consent cannot be given by or on behalf of the data subject, or (ii) the data controller cannot reasonably be expected to obtain the consent of the data subject, or (b)in order to protect the vital interests of another person, where consent has been unreasonably withheld.
- The processing is carried out in the course of its legitimate activities of a not for profit body and relates only to members or relevant contacts and the body does not disclose without consent
- The information contained in the personal data has been made public as a result of steps deliberately taken by the data subject.
- The processing is necessary for medical purposes including for preventative medicine, medical diagnosis, medical research, the provision of care and treatment and the management of healthcare services.
- The processing is for racial or ethnic equal opportunities purposes

2 Personal data shall be obtained **only for one or more specified and lawful purposes**, and shall not be further processed in any manner incompatible with that purpose or those purposes.

3 Personal data shall be **adequate, relevant and not excessive** in relation to the purpose or purposes for which they are processed.

4 Personal data shall be **accurate** and, where necessary, kept **up to date.**

5 Personal data processed for any purpose or purposes shall **not be kept for longer than is necessary** for that purpose or those purposes.

6 Personal data shall be processed **in accordance with the rights of data subjects** under this Act.

7 Appropriate technical and organisational **measures** shall be taken against unauthorised or unlawful processing of personal data and **against accidental loss or destruction of, or damage to, personal data**.

8 Personal data **shall not be transferred to a country or territory outside the European Economic Area** unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.

SHIP will be concerned with data for research purposes only so the exemption for "Research, History and Statistics" in section 33 will apply. This provides that the data can be kept indefinitely if they are processed in compliance with the relevant conditions[5], and allows exemption from section 7 which pertains to principle 6.

DPA Principle	**How SHIP Addresses the Principle**
1: Fair and lawful processing	All data will be processed within the act. Sensitive personal data eg health data, can legitimately be processed without consent if it is necessary in pursuance of legitimate interests of the data controller or by the third party or parties to whom the data are disclosed, provided it does not prejudice the rights and freedoms or legitimate interests of the data subject (schedule 2) and where processing is for medical research (schedule 3).
2: Specified lawful purposes	Application form will detail the research question and permissions will be specific to that purpose. All downstream events will be in fulfilment of that purpose.
3: Adequate relevant and not excessive	Researchers will only be supplied with the data they need for their stated research question

[5] (a) that the data are not processed to support measures or decisions with respect to particular individuals, and
(b) that the data are not processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject.

4: Accuracy of data	Accuracy of data accessed via SHIP is the responsibility of the data controller of the source data. SHIP will not alter the data except to obscure identity of data subjects when the agreement of data controller and researcher will be obtained. Datasets accessed via SHIP will generally be considered as "snapshots" so no attempt will be made to update them in real time. Longitudinal datasets will take the form of a series of snapshots.
5: Not longer than necessary	Under section 33 this need not apply because the relevant conditions will be met i.e. (a) that the data are not processed to support measures or decisions with respect to particular individuals, and (b) that the data are not processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject.
6: In accordance with rights	Under section 33 this need not apply because the relevant conditions will be met (see above).
7: Measures against loss or destruction	Measures against data loss or destruction are detailed in the System Security Policy for SHIS-r, a document of NHS Information Services Division written by Anthea Springbett.
8: Not transferred beyond EEA	SHIP will not transfer data to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data