A Charter for Safe Havens in Scotland

Handling Unconsented Data from National Health Service Patient Records to Support Research and Statistics

November 2015

The Scottish Government
Riaghaltas na h-Alba

Purpose

The information held in electronic patient health records offers enormous opportunities for research that can lead to more effective, safer health services and treatments, and better understanding and new insights into the causes and development of diseases. Whilst the public are generally very supportive of the aims of health informatics research (research using electronic health records), there is significant public concern about how the data in their health records might be used, who has access to these data, and that as a result of this research their privacy and confidentiality may be compromised.

In order to address these concerns and support health informatics research, a robust governance framework has been developed to ensure that health informatics research in Scotland is ethical, in the public interest, scientifically sound, and patient identity and privacy are appropriately protected. 'Safe Havens' are a crucial element of this framework in order to protect identity and privacy and facilitate health informatics research.

This charter sets out the agreed principles and standards for the routine operation of Safe Havens in Scotland where data from electronic National Health Service (NHS) patient records can be processed, linked with other data and analysed to support research when it is not practicable to obtain individual patient consent while protecting patient identity and privacy. It also describes, at a high level, how Safe Havens will work together across Scotland on collaborative research projects as part of a federated network.

Background

Health research can be conducted very efficiently and effectively through the use and linkage of routinely collected data held in electronic patient records. Many important health research studies may only be amenable practically using this approach. Through health informatics research, new knowledge can lead to improvements in health by, for example, understanding better the causes of disease, the effectiveness of drugs, or the impact of health services[1]. However, alongside the potential benefits and opportunities, there are significant public concerns that patient privacy and confidentiality could be compromised through the use of these data and/or the data could be misused.

In order to address these concerns, health informatics research must be well-controlled and encompass robust governance processes that ensure NHS data are

[1] Health Informatics Research Advisory Group (2015). A Health and Biomedical Informatics Research Strategy for Scotland; Enhancing research capability in health informatics for patient and public benefit. *http://www.gov.scot/Resource/0047/00475145.pdf*

only used for approved purposes and to safeguard patient identity and privacy[2,3,4]. The governance processes must be such that they provide assurance to patients, the public, and NHS organisations that hold the source data and have legal responsibilities for protecting them, that data from electronic patient records can be used safely and in a trustworthy way without compromising patient identity and privacy.

The establishment of Safe Havens has been acknowledged as a means by which robust controls and safeguards can be put in place[1,2,3,4]. Safe Havens operating under an accreditation framework have been considered to be the most appropriate environments to facilitate research using de-identified (also termed 'pseudonymised') data from electronic patient records when it is not practicable to obtain specific consent from the individuals for the use of data in their records[4].

Safe Havens are specialised, secure environments supported by trained, specialist staff where data in electronic patient records can be processed and linked with other health data (and/or non-health-related data) and made available for analysis to facilitate research while protecting patient identity and privacy. Risk of identification is minimised through 'pseudonymisation' of the data, stripping away information that is not required for the research study and information that would allow individuals to be identified directly (for example, names, addresses), and also through the robust safeguards in operation at the Safe Haven to protect patient identity. These safeguards include the separation of the indexing/linking platforms, where de-identified data are linked, and the analytical platforms where the newly created linked de-identified datasets can then be analysed (see Figure 1). The safeguards also include the use of agreed formal standard operating procedures by the Safe Haven support staff with compliance monitored. Furthermore, access to the data in the Safe Haven is tightly controlled; only approved and vetted researchers are permitted access to undertake analyses and the source data are never released from the Safe Haven. These safeguards are described in this charter.

[2] Thomas & Walport (2008). Data Sharing Review. http://webarchive.nationalarchives.gov.uk/+/http:/www.justice.gov.uk/docs/data-sharing-review.pdf
[3] UK Administrative Data Research Network (2012). Improving Access for Research and Policy. http://www.esrc.ac.uk/files/publications/themed-publications/improving-access-for-research-and-policy/
[4] Caldicott (2013). Information: to Share or Not to Share? The Information Governance Review. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf

The Safe Havens in Scotland that handle data from NHS patient records for research operate within a robust research governance framework. Firstly, they can only receive and process data under the express agreement of Data Controllers[5] – the NHS or other organisations holding the source records which have legal obligations to protect the data within them – and when the Data Controllers are satisfied about the safeguards Safe Havens have in place to protect patient identity. Secondly, before research projects using data from electronic patient records can begin, they are considered by expert ethics and scientific panels in a similar way to studies involving patients. Projects are also scrutinised by an expert panel, such as the Public Benefits and Privacy Panel (PBPP)[6] or the local Caldicott Guardians acting on behalf of the Data Controllers. The panel assesses the public benefit of the research study against the risk that individual privacy might be compromised, and also that any information releases are carefully controlled. Through these processes careful consideration is given to whether specific health informatics research studies should be done, and whether the safeguards in place provide appropriate protection of the identity and privacy of patients since they themselves cannot, for practical reasons, be approached for their specific consent. Only when the conditions of the panels are met can a research study proceed. Any decision to provide access to data must follow NHSS scrutiny, when involving NHS data, and the use of an independently accredited safe haven provides additional assurance.

The Charter for Safe Havens in Scotland

This charter sets out the principles and standards for the routine operation of the Safe Havens in Scotland that are processing and linking data from electronic NHS patient records when it is not practicable to obtain individual patient consent. It draws heavily on other documents in particular: *the Guiding Principles for Data Linkage*[7] (which in turn draws on: Human Rights Legislation, the Data Protection Act, Guidance from the Information Commissioner, and the Scottish Government Identity Management and Privacy Principles), the *SHIP Blueprint*[8]and associated governance framework that defines standards and process for the use of non-consented linked data for health informatics research in Scotland[9].

[5] Data Controller should be taken to mean the individual NHS board or organisation with responsibilities as a Data Controller, as defined in the Data Protection Act 1998, and registered with the Information Commissioner's Office, agents acting on their behalf, such as individual Caldicott Guardians, and/or governance or scrutiny mechanisms operating with delegated decision making on their behalf.
[6] The Public Benefit and Privacy Panel for Health and Social Care http://www.informationgovernance.scot.nhs.uk/
[7] Scottish Government (2012). Joined up data for better decisions: Guiding Principles for Data Linkage. http://www.scotland.gov.uk/Resource/0040/00407739.pdf
[8] SHIP Blueprint (2012). http://www.scot-ship.ac.uk/publications.html
[9] SHIP Guiding Principles and Best Practices (2010). http://www.scot-ship.ac.uk/sites/default/files/Reports/Guiding_Principles_and_Best_Practices_221010.pdf

Through the adoption of this charter the governance around Safe Havens will be enhanced and this charter will be supported through a programme of independent accreditation of Safe Havens. An accreditation process is under development. All Safe Havens handling data from NHS patient records to support research must adhere to this charter and, once a formal accreditation process has been established, they must be accredited in order to operate. Accreditation, once established, will be a key criterion for privacy panel assessment and approval of research projects and may be a requirement for Scottish Government funding.

There may be instances when, in order to support an important research study, one or more Safe Havens may need to depart from aspects of this charter. Under these circumstances the Safe Haven(s) will be required to seek specific additional permissions and approvals including from Data Controller(s) so that additional assurance is in place before the research study starts. Release of data may also be approved when Data Controllers (or their delegated authority, such as the PBPP) considering the balance of benefits and risks are reassured that appropriate safeguards are in place to protect the privacy of individuals, and release is specified explicitly within the data sharing agreement with the Data Controllers.

This charter has been developed in consultation with the Safe Havens created by NHS Scotland Boards, Caldicott Guardians, NHS Research and Development Directors, University-based researchers and Scottish Government experts in data governance. This is an evolving area, therefore, the charter will be kept under review and will be reviewed within two years of its introduction to ensure that it is meeting its aims, takes into account technical developments, supports proportionate governance and aligns with other developments within Scotland and across the UK and internationally.

The charter also sets out the relationship, at a high level, of a federated network of Safe Havens in Scotland that are using data from NHS patient records.

A Federated Network of Safe Havens

Creation of Safe Havens is expensive, technically challenging and highly specialised in terms of the technological and governance requirements that need to be in place. Hence NHS Boards have created Safe Havens in conjunction with Universities with specialised units or with accredited commercial organisations who employ core support staff who provide a managed technical Safe Haven service as data processors. These services can be offered to other NHS Boards under robust governance arrangements agreed with the Data Controllers of the contracting Boards. Some units also offer other informatics services in addition to the Safe Haven service.

Currently, there are five Safe Haven services in Scotland that have been commissioned by NHS Scotland Boards: one in each of the four lead regional NHS Research Scotland (NRS) nodes (in Aberdeen, Dundee, Edinburgh and Glasgow),

and a National Safe Haven within Information Services Division (ISD), an expert unit within the Common Services Agency (known as NHS National Services Scotland). . The National Safe Haven is part of the Scottish Informatics Linkage Collaboration (SILC) that also includes the electronic Data Research and Innovation Service (eDRIS)[10] and the National Records of Scotland indexing service. SILC facilitates linkage for research and statistical activities across many sectors including the NHS. The initial technical infrastructure that supports SILC has been funded by the MRC, Scottish Government and NSS through the Farr Institute Scotland[11], a collaboration between six Scottish Universities and NHS National Services Scotland. The NRS nodes received funding, through NRS infrastructure allocations from the Chief Scientist Office, to help establish Safe Havens.

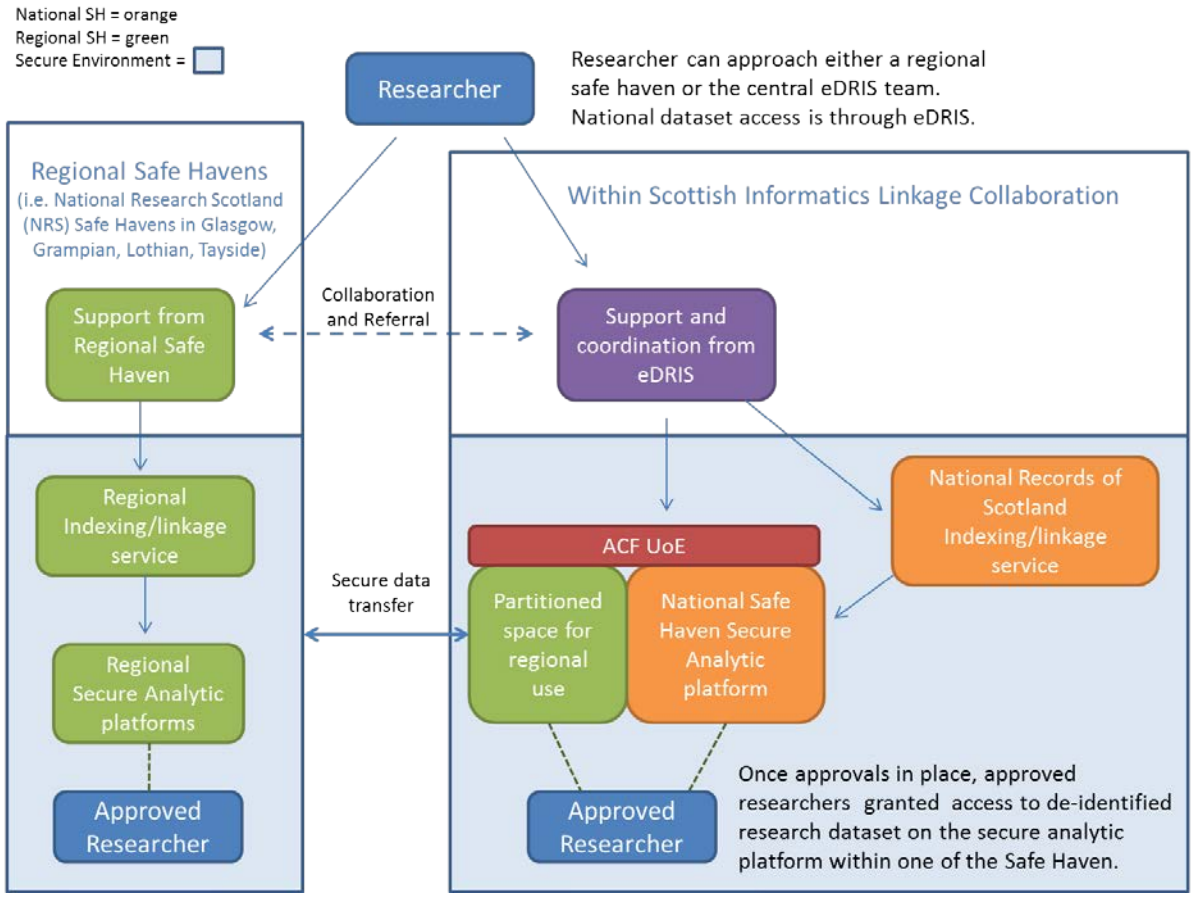

Figure 1: The federated network of Safe Havens created by NHS Scotland Boards. eDRIS=electronic Data and Research Innovation Service, ACF UoE=Advanced Computing Facility University of Edinburgh.

Together, the National Safe Haven within SILC and the four NRS Safe Havens have formed a federated network of Safe Havens in order to work collaboratively to support health informatics research across Scotland (see Figure 1). All the Safe

[10] http://www.isdscotland.org/Products-and-Services/EDRIS/
[11] http://www.farrinstitute.org/centre/Scotland/3_About.html

Havens have individual responsibility to operate at all times in full compliance with all relevant codes of practice, legislation, statutory orders and in accordance with current good professional practice. Each Safe Haven may also work independently to provide advice and assistance to researchers as well as secure environments to enable health informatics research on the pseudonomyised research datasets they create. Co-ordination and close collaboration across Scotland is key for international competitiveness; the establishment of this charter and the network will facilitate collaboration between the Safe Havens by ensuring that they all work to the same principles and standards. The governance of SILC involves representatives of the Safe Havens. The technical operation of the federated network and the integration of the Safe Havens with SILC is set out elsewhere[12].

eDRIS was established as a specific ISD function within NSS and provides a single point of contact for advice on research project design and development as well as access via the national Safe Haven to a wide range of national datasets. Given the well-established close working relationship with ISD (the Data Controller of national NHS Scotland datasets), the National Safe Haven may be best placed to take the lead when research requires the processing and linkage of national datasets. Similarly, it is anticipated that data controllers in NHS Boards will select the regional Safe Havens in the NRS nodes to support research that involves the processing and linkage of data using local/regional NHS datasets[6]. However, the federated network allows Safe Havens to work in collaboration, for example, when research studies may require datasets from more than one NHS region to be combined and/or where the specialist skills of a number of Safe Havens can be applied to facilitate a study that requires a range of expertise[6].

eDRIS can act as a first port of call for researchers on behalf of the federated network to coordinate research and advise researchers about what the Safe Havens in the federated network can provide, although each Safe Haven can also be approached by, and work with, researchers directly on regional Health Board data.

All the Safe Havens are committed to the principles and standards set out in this charter and to undergo formal accreditation once an independent accreditation process has been established. The compliance of the Safe Havens to the principles and standards set out in the charter will be assessed regularly by the independent accrediting body and accreditation will be a requirement to operate, once the accreditation system has been established.

Should new Safe Havens be commissioned by NHS Scotland Boards, they will need to agree to operate under this charter, undergo accreditation, and comply with the technical specifications required to link securely to the Safe Havens within the federated network to be a part of the network.

[12] The operation of SILC and the technical operation of the federated network will be published elsewhere.

Principles for the operation of Safe Havens in Scotland handling data from NHS patient records

The principles for the routine operation of the Safe Havens handling data from NHS patient records are set out below. For each principle, the principle (in bold), the background to the principle (in italics), and (at a high level) how Safe Havens implement each principle (in normal text) are described. Operating standards and procedures that set out in more detail how the principles will be achieved in practice are given in the Technical Annex.

In this paper, reference to Data Controller should be taken to mean the individual NHS board or organisation with responsibilities as a Data Controller, as defined in the Data Protection Act 1998, and registered with the Information Commissioner's Office, agents acting on their behalf, such as individual Caldicott Guardians, and/or governance or scrutiny mechanisms operating delegated decision making on their behalf.

Principle 1

Safe Havens will provide secure environments that are trusted by NHS Data Controllers which allow the safe and secure transfer of data with maximum fidelity from Data Controllers in Health Boards (and where applicable other bodies) to Safe Havens and between Safe Havens.

The Guiding Principles for Data Linkage states that:

"Security of data transfer, storage and use is vital for the protection of privacy, especially where there is any risk of reidentification."

"Appropriate and proportionate physical and technical security measures should be applied to ensure the confidentiality, integrity and availability of information and should reflect the assessment risk level of information assets."

"The default position should be that data users have access only to data from which names and direct identifiers have been removed, and data users should be subject to obligations not to attempt to re-identify individual data subjects. Any requirement for researchers to have access to data containing identifiers should be fully justified and risk assessed."

Safe havens will ensure that data are only transferred and held within secure networks (such as the NHS N3 or SWAN networks) or using secure file transfer protocols that encrypt data in flight to ensure that individuals' privacy is protected.

Research datasets will be held on electronically Secure Analytic Platforms in physically secure data centres, with access provided either from a 'Secure Safe Setting' (i.e. a specified and sanctioned physical location) or via a Virtual Private Network or encrypted communication sessions. Data Controllers will be responsible

for determining the appropriate route of access. Researchers will not be able to add or remove any information from the Secure Analytic Platform before it has been reviewed through methods applied by the Safe Haven to ensure that individuals' privacy is protected.

Principle 2

Safe Havens must not independently develop nor retain non-consented datasets or linked datasets that could potentially be used to identify individuals unless the development, use and retention of these datasets are described clearly and set out in the data sharing agreements with the Data Controller(s) in the Health Board(s) (and where applicable other bodies) that hold the source records.

The Guiding Principles for Data Linkage states that:

"Linked datasets should be kept for the minimal time necessary for the original purpose of the linkage to be met. The onus is on those wishing to hold datasets for longer to justify this, e.g. by demonstrating that adequate anonymisation takes the data outside the remit of the data protection regime. If a secondary purpose arises, a new Privacy Impact Assessment should be considered, and data-sharing agreements revised."

Safe Havens will not act as data repositories to collate, maintain, curate and use datasets or linked datasets of potentially identifiable data without the explicit agreement of the Data Controller(s) in the Health Board(s) (and where applicable other bodies) holding the source records, and unless specific provisions are made and set out in the data sharing agreements between the Safe Havens and the Data Controller(s).

Principle 3

As Safe Havens will act as Data Processor(s) on behalf of the Data Controller(s) of NHS Scotland Board(s), an application to become a Safe Haven which can safely and securely process NHS data must be approved by the appropriate Health Board Chief Executive(s) or the delegated authorities within the Health Board(s) (for example the Caldicott Guardian). Accreditation, once established, can facilitate continued approval.

The Guiding Principles for Data Linkage states that:

"Every reasonable effort should be made to consider and minimise the risks of identification (or re-identification) to data subjects and their families arising from all aspects of data handling.

"Where obtaining consent is not practicable, then removal of direct identifiers should occur as soon as is reasonably practicable…"

"Procedures to link data should involve the separation of identifiers (e.g. name, or unique reference number) from the rest of the data, and consideration should be given to separating the indexing, linking and analysis functions and personnel."

"The linkage method used should be that which requires the minimum necessary identifiable data."

"Data controllers should determine and agree upon the appropriate extent of anonymisation to be applied to any given dataset or linkage exercise."

In order to obtain the approval from NHS Board(s), Safe Havens need to ensure that all data are held securely and in accordance with the instructions of the Data Controller(s) holding the source records, and that they will process data only in ways approved by the Data Controller(s). Data within analytic platforms of Safe Havens must not contain personal identifiers (for example names, addresses etc.). The Safe Haven will need to demonstrate that there are robust processes in place for de-identification so that data are pseudonymised before entering the analytic platform of the Safe Haven for analysis. Safe Haven accreditation, once established, will provide a mechanism for ensuring robust procedures and processes are in place.

Principle 4

All staff working within Safe Havens who are providing the Safe Haven service will be trained in Information Governance and the law relating to the protection of individuals' privacy (such as the Data Protection Act) and will be trained on and work to written standard operating procedures. Both staff and operating procedures will be subject to monitoring as well as regular review and audit.

The Guiding Principles for Data Linkage states that:

"All personnel involved in data linkage activities should be properly trained on the data security policies and procedures, and should undertake periodic refresher training.

"The importance of data security should be reflected in the business objectives of all organisations involved in data linkage."

"Information about data security policies and procedures should be highly visible within organisations conducting indexing or linking or sharing of personal data."

"All practices, including all data linkages, shall be appropriately monitored and regulated by a relevant individual, organisation or governance body."

Staff involved in providing the Safe Haven service must be bound through contractual requirements to protect individuals' privacy and be subject to sanctions if they fail to fulfil these requirements. The Safe Haven must ensure that these staff are trained on and work to written standard operating procedures with regular

internal monitoring, review and audit of procedures and their operation. Audit records must be kept of staff members' access to personal identifiable information.

Principle 5

Safe Havens work in partnership with academia, public service providers and industry to undertake research using de-identified or anonymised data that is in the public interest. However, personal data cannot be sold by a Safe Haven or transferred to a commercial organisation. Nor can they be transferred, nor access provided, to a third party (i.e. researchers or others) unless specified explicitly by the Data Controller(s) holding the source records and unless the third party operates to, at minimum, equivalent standards and with equivalent safeguards.

Findings from research on public attitudes suggests that, whilst the use of anonymised patient data for publicly funded research is generally accepted, attitudes to commercial use of patient data are much more ambivalent and are dependent on the research aims and whether or not public benefits are likely[13,14,15].

The Guiding Principles for Data Linkage states that:

"Benefits arising from linkage of personal data are public goods and should be shared as widely as possible."

"Where linkages resulting in commercial gain are envisaged, this should be clearly and publicly articulated and widely communicated."

Safe Havens must not sell personal data nor transfer sensitive personal information (or information which is likely to allow the identification of individuals) to commercial organisations. Nevertheless, Safe Havens should promote cross-sector partnership working between industry, academia and public service providers to undertake research using de-identified data that is in the interests of the people of Scotland, through open and transparent managed collaborations. However, as with all access to linked datasets for research, the data cannot be copied, nor removed from the Safe Haven, nor can they be released outside of the Safe Haven (unless shared between Safe Havens within the federated network) unless specified explicitly within the data sharing agreement with the Data Controller(s) holding the source records

[13] Davidson *et al.* (2012) Public acceptability of cross-sectoral data linkage. Deliberative research findings. Scottish Government. http://www.scotland.gov.uk/Publications/2012/08/9455
[14] Davidson *et al.* (2013) Public acceptability of data sharing between the public, private and third sectors for research purposes. Scottish Government. http://www.scotland.gov.uk/Publications/2013/10/1304
[15] Armstrong *et al.* Public attitudes to research governance. A qualitative study in a deliberative context. Wellcome Trust. http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtx038443.pdf

and unless the third party operates with, at least, equivalent standards and safeguards.

Principle 6

Only researchers who have completed nationally approved safe researcher training satisfactorily and have been vetted by a designated member of the Safe Haven support staff will have access to research datasets in or through that Safe Haven.

The Guiding Principles for Data Linkage states that:

"All data recipients should be appropriately vetted to ensure they have adequate training. Vetting procedures should be robust and transparent and proportionate to the requests made and the sensitivity of the data requested."

"The terms and conditions for data sharing should be set out in the form of a data sharing agreement. Where researchers wish to deviate from or modify the terms of the data use/sharing agreement, new terms must be agreed by all parties."

Only researchers (this may include supervised students for the purposes of training) that have completed a nationally approved and/or validated course, once these have become established, which covers information governance, privacy protection and the relevant legislation, can access research datasets in Safe Havens. Approved researchers will work for a recognised academic, public sector or industry organisation, have agreed to a 'terms of use policy' and will be liable to sanctions should they breach the terms of use. Approved researchers from industry will access research databases as part of managed collaborations with recognised academic or public sector organisations. Sanctions can be applied at individual, research group, and institutional levels. Approved status is verified by the Safe Haven before each time of granting access to the Safe Haven and a single National Register of Approved Researchers will be kept and maintained by eDRIS on behalf of the federated network.

Principle 7

Organisations hosting Safe Havens within the federated network of Safe Havens in Scotland will participate in the development of the governance and research activities of data linkage across Scotland and the collaborative working across the federated network.

Through the collaborative opportunities provided through the federated network, best practice will be shared and the federated network of Safe Havens in Scotland will function as a single research site when it makes sense to do so. Appropriate governance mechanisms will be established to support this activity.

Technical Annex

Operating standards and procedures for Safe Havens in Scotland to enable the Principles set out in this Charter and to address data security and privacy requirements are set out below. Each Safe Haven must be supported by a designated senior professional who is responsible for the operation of the Safe Haven.

1. Safe Havens function as a Data Processor for any given dataset, agree a mandate with each Data Controller to ensure activity is centrally logged, monitored and audited, and act only in accordance with the explicit instructions from each Data Controller. Established national or local data privacy and scrutiny bodies comprising appropriate expert and lay members can make assessments on behalf of Data Controllers about the risks and benefits of data releases to Safe Havens which must then operate in strict accordance with their specific mandate.

2. Where a Data Controller provides data to a Safe Haven located within their organisation:

    - the staff providing the data to the Safe Haven, and the Safe Haven staff should be in separate management units and accountable to different line managers to minimise conflicts of interest arising within these roles; and

    - other than where agreed explicitly for purposes of data sensitivity or quality, linkage and analysis should be undertaken by individuals in different roles

    - the Safe Haven staff must comply with the instructions and mandate agreed with the Data Controller.

3. Safe Havens must maintain accurate records of:

    - all policies and written agreements underpinning the operation of the Safe Haven

    - the names, roles and levels of permissions to view and process data of all staff employed within the Safe Haven

    - the names and roles of all those staff given access to data, alongside summary information of the data accessed and the purpose for which access was approved

    - all projects conducted or supported through provision of data, information about who approved the project and a summary of the analytical outputs

    - data received into the Safe Haven and review date and deletion date if applicable

- cross reference with Caldicott approval, IRAS registration of the dataset, research registration
- data sharing agreements
- collaboration agreements
- inspection and other regulatory reports
- release of aggregated data in pursuit of open data agreements
- publications

4. Safe Havens must ensure all Safe Haven staff undertake training which addresses Information Governance and the relevant data protection legislation and regular refresher training as required.
5. Safe Havens must include confidentiality clauses within the contractual conditions of all staff involved in the management, processing or use of data, and instigate disciplinary procedures in the event of contractual conditions being breached.
6. Safe Havens must hold and process all de-identified data and potentially identifiable data exclusively and separately within restricted access areas within secure networks[16].
7. Systems should comply with relevant ISO standards. Oversight of systems security and compliance should be the responsibility of a designated security officer.
8. Safe havens must conduct penetration testing every two years; both from outwith and within the Safe Haven environment.
9. Safe Havens must restrict physical access to any room within which identifiable or potentially-identifiable data are stored in paper form.
10. Safe Havens must restrict physical access to any room within which the servers hosting identifiable or potentially-identifiable data electronically are stored.
11. Safe Havens must receive and transfer data only when necessary and do so within a secure network (NHS N3, the Scottish Wide Area Network (SWAN) or a network with equivalent controls for comparable data). Where use of a secure network is not possible, a secure method for file transfer must be used, such as Secure File Transfer Protocol (SFTP).
12. Safe Havens must develop a publication plan and publish a list of all active data sharing agreements on their websites to increase public understanding of data

[16] Within the NHS this is the NHS N3 Network or equivalent and processed in line with NHS security standards and Health Board policy.

use, and to ensure information on data sources is accessible and discoverable, so that potential users can also find out about data resources and how to apply for access.

In the creation of project specific datasets from a single data controller Safe Havens must:

13. Remove all direct identifying information and replace them with a project specific unique identifier.

14. Retain project datasets (data extracts or linked datasets) in an analytic environment for the time period specified through written agreement with the Data Controller(s) and subsequently archive or delete.

15. Archive data in a secure environment for a specified period of time only in accordance with the specific written agreement with the Data Controller(s). Archived data must not be accessed for any purpose other than the original research unless by written agreement with the Data Controller(s). Clear and transparent records of archived data, review and planned deletion dates must be maintained.

16. Disclosure assessment and disclosure control will be applied before data are provided to an Approved Researcher in an Analytic Platform.

In the creation of project specific linked datasets from multiple data controllers Safe Havens must:

17. In addition to 13 to 16, undertake data linkage in a manner that separates the functions of the indexer/linker and researcher with the objective of minimising the number of staff with access to identifiable information. A written description of how this standard is complied with should be recorded for each linkage.

In delivering the function of a <u>Secure Analytics Platform</u> Safe Havens must:

18. Ensure project data are only placed into the Secure Analytics Platform by the designated staff that provide the support for the Safe Haven.

19. Only allow Approved Researchers access to data on written instruction from the Data Controller and with strict adherence to all conditions laid down in relevant Data Governance documentation (e.g. data sharing agreements, user agreements etc.).

20. Permit access to data only to Approved Researchers and via two factor authentication log-in.

21. Permit remote access to data via a Virtual Private Network or using encrypted communication sessions only with the agreement of the Data Controller(s),

otherwise permit access only via a secure physical terminal within a secure Safe Haven room.

22. Never allow Approved Researchers access to direct identifiers without direct written instruction from the Data Controller. Such instructions should not be part of the standard approach: under most circumstances only project specific unique identifiers should be accessible to the Researcher.

23. Minimise the risk of study data being copied or removed from the Secure Analytics Platform by an Approved Researcher.

24. Allow analytical outputs (e.g. reports, summaries, aggregate statistics, graphs etc.) to be downloaded only after they have been checked for statistical disclosure by designated analytical staff supporting the Safe Haven if instructed by the Data Controller(s).

25. Retain, for governance purposes, copies of all analytical outputs which leave the Analytic Platform.

In delivering the function of a <u>Secure Analytics Platform</u> Safe Havens should:

26. In conjunction with 21, provide Approved Researchers with a view of the study specific dataset via a secure remote-access environment (e.g. Citrix) to enable remote access while mitigating the risk of data being removed from the Secure Analytics Platform without permission and minimise the risk of the introduction of viruses or malware to the analytic environment.

27. Facilitate the uploading of user-specific analytic files (e.g. look-up tables, statistics scripts) or bespoke applications with careful risk assessment and consideration of how to minimise the risk of the introduction of viruses or malware to the analytic environment.

<u>Secure Safe Setting</u>

28. Safe Havens should provide or sanction access points that meet the requirements of a Secure Safe Setting and allow researchers to access data held in any of the Secure Analytics Platforms across Scotland via a ‘thin client’ mechanism (assuming appropriate permissions are in place).

29. Secure Safe Settings consist of ‘thin client’ terminals that are located in secure physical environments where the researchers’ behaviours and actions are monitored. An audit log of who has accessed which data should be kept.

Glossary

Anonymised data

Data in a form that does not identify individuals and where re-identification is not possible routinely.

Approved Researcher

A professional researcher who satisfies the conditions of the relevant body (e.g. eDRIS) regarding qualifications, Data Protection and Information Governance training, affiliation and contract.

Data Controller

An individual, organisation or body that determines the purposes for which and the manner in which any personal data are, or are to be, processed. An annex to the Guiding Principles explains the roles and responsibilities of data controllers. See www.scotland.gov.uk/GuidingPrinciplesforDataLinkage

Data processor

A person who processes data on their own behalf or on instruction from the Data Controller but who does not determine the purpose and manner in which the data are processed. For the purposes of this charter the data processor is the Safe Haven and all the staff involved in providing this service.

Data Linkage / Record Linkage

Data linkage is the joining of two or more datasets using individual reference numbers / identifiers or statistical methods of matching such as probabilistic modelling.

De-identification (pseudonymisation)

The process of distinguishing individuals in a dataset by using a unique identifier (code) which does not reveal their 'real world' identity.

electronic Data Research and Innovation Service (eDRIS)

The electronic Data Research and Innovation Service (eDRIS) eDRIS is part of Information Services Division within NSS. It provides a single point of contact to assist researchers in study design, approvals and data access in a secure environment.

Farr Institute

The Farr Institute is a UK collaboration to harness health data for patient and public benefit by setting the international standard for the safe and secure use of electronic patient records and other population-based datasets for research purposes. There are four centres across the UK, of which the Farr Institute-Scotland is one.

Health informatics research

Research using the data from electronic patient records. It usually involves combining (linking) data from a number of different sources.

Individual Reference / Identifier

Frequently a sequence of characters and/or numbers that is used and/or assigned by an organisation to a person to identify uniquely the person for the purposes of the organisation's systems and operations.

Information Services Division (ISD)

Information Services Division (ISD) is part of NHS National Services Scotland and provides statistical information and analysis for the NHS.

National Records of Scotland

The National Records of Scotland performs the registration and statistical functions for the Registrar General for Scotland, including responsibility demographic statistics and census and archival functions.

NHS National Services Scotland (NSS)

NHS National Services Scotland (NSS) is the organisation which provides advice and services to the rest of NHS Scotland. It includes Information Services Division.

NHS Research Scotland (NRS)

NHS Research Scotland (NRS) is a partnership involving Scottish NHS Boards and the Chief Scientist Office (CSO) of the Scottish Government with the aim of ensuring that NHS Scotland provides the best environment to support clinical research.

Personal Data / Identifiable Data / Personal Information

Information about an individual who can be identified from that information with or without other information.

Pseudonymisation (de-identification)

The process of distinguishing individuals in a dataset by using a unique identifier (code) which does not reveal their 'real world' identity.

Public Benefit and Privacy Panel for Health and Social Care

Governance structure of NHSScotland established with delegated authority from NHSScotland Chief Executive Officers and the Registrar General to scrutinise any use of NHSScotland controlled data and NHS Central Register data.

The http://www.informationgovernance.scot.nhs.uk/

Re-identification

The process of turning de-identified data back into personal data through the use of data matching or similar techniques.

Safe Haven

A term used to explain either a secure physical or remotely accessed environment, supported by trained specialist staff working under an agreed set of administrative arrangements with an organisation to ensure confidential personal information is processed, analysed and/or communicated safely and securely.

Scottish Informatics Linkage Collaboration (SILC)

Scottish Informatics Linkage Collaboration (SILC) is a collaboration between academic and public bodies in Scotland to deliver shared services and ensure that Scotland realises the benefits that can be derived through the legal, ethical and carefully controlled use of administrative, survey and other types of data.

Further Reading

R Thomas and M Walport (2008) Data Sharing Review
http://webarchive.nationalarchives.gov.uk/+/http:/www.justice.gov.uk/docs/data-sharing-review.pdf

Information to share or not to share? Information Governance Review (2013)
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf

Anonymisation Code of Practice
http://ico.org.uk/for_organisations/data_protection/topic_guides/anonymisation

Data sharing code of practice
https://ico.org.uk/media/about-the-ico/consultations/2069/data_sharing_code_of_practice.pdf

Data Sharing: Legal Guidance for the Public Sector
http://www.scotland.gov.uk/Publications/2004/10/20158/45784

Identifying 'data controllers' and 'data processors' Data Protection Act 1998
http://ico.org.uk/~/media/documents/library/Data_Protection/Detailed_specialist_guides/data-controllers-and-data-processors-dp-guidance.pdf

Identity Management and Privacy Principles
http://www.scotland.gov.uk/Topics/Government/PublicServiceReform/efficientgovernment/privacyprinciples

Joined up data for better decisions: Guiding Principles for Data Linkage
http://www.scotland.gov.uk/Publications/2012/11/9015

Christie Commission on the Future Delivery of Public services
http://www.scotland.gov.uk/Resource/Doc/352649/0118638.pdf

This publication is available at **www.gov.scot**

Any enquiries regarding this publication should be sent to us at
The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

ISBN: 978-1-78544-496-8 (web only)

Published by The Scottish Government, November 2015

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA
PPDAS45015 (11/15)