The Information Centre
knowledge for care

NHS
Connecting for Health

Pseudonymisation Implementation Project (PIP)

Implementation Guidance on Local NHS Data Usage and Governance for Secondary Uses

Final v1.0 - 20 November 2009

© Crown Copyright 2009 Without Prejudice

PIP Implementation Guidance			
Programme	NPFIT	Document Record ID Key	
Sub-Prog / Project	Pseudonymisation Implementation Project (PIP)	**NPFIT-FNT-TO-BPR-0022.01**	
Prog. Director	J Thorp	Version	01
Owner	.	Status	Final
Author	Wally Gowing	Version Date	20 November 2009

Document Status:

This is a controlled document.
Whilst this document may be printed, the electronic version maintained in FileCM is the controlled copy. Any printed copies of the document are not controlled.

Related Documents:

These documents will provide additional information.

Ref	Doc Reference Number	Title	Version
	NPFIT-FNT-TO-BPR-0022.01	PIP Implementation Guidance	FV1
1	NPFIT-FNT-TO-BPR-0023.01	Reference Paper 1 - Terminology[1]	FV1
2	NPFIT-FNT-TO-BPR-0024.01	Reference Paper 2 – Business Processes and New Safe Havens[2]	FV1
3	NPFIT-FNT-TO-BPR-0025.01	Reference Paper 3 – De-identification[3]	FV1
4	TBA	Reference Paper 4 – Technical White Paper[4]	FV1
5	dh_4069254	NHS Code of Practice on Confidentiality[5]	
6	NA	PIP Planning Template and Guidance[6]	

[1] **http://www.connectingforhealth.nhs.uk/systemsandservices/sus/delivery/pseudo**
[2] **http://www.connectingforhealth.nhs.uk/systemsandservices/sus/delivery/pseudo**
[3] **http://www.connectingforhealth.nhs.uk/systemsandservices/sus/delivery/pseudo**
[4] **http://www.connectingforhealth.nhs.uk/systemsandservices/sus/delivery/pseudo**
[5] [5]

www.dh.gov.uk/en/Managingyourorganisation/Informationpolicy/Patientconfidentialityandcaldicottguardians/DH_4100550
[6] **http://www.connectingforhealth.nhs.uk/systemsandservices/sus/delivery/pseudo**

© Crown Copyright 2009

Contents

© Crown Copyright 2009

Pseudonymisation Implementation Project Implementation Guidance

1 Introduction

1.1 Purpose and scope

This paper provides the local implementation guidance for the Pseudonymisation Implementation Project (PIP) and links to other PIP documents that provide reference material, guidance and general information.

The purpose of this paper is to:

- Provide a pathway through the steps that need to be taken to implement local data usage and governance from the viewpoint of different organisations involved in running NHS business operations
- Provide high-level guidance on governance arrangements to ensure that 'effective anonymisation' can be achieved to mitigate risk associated with use of patient data.
- Provide high-level guidance on governance arrangements and technical solutions for de-identification, which will ensure the security and confidentiality of person level data within local systems, referred to within PIP as implementing local data usage and governance
- Provide links to the separate elements of more specific guidance covering terminology, new safe havens and business processes, de-identification and security.
- The scope of the document is therefore concerned with the management and use of patient identifiable in relation to NHS business operations; it is not concerned with the use of such data for medical research purposes.

1.2 PIP Aims and Implementation

The overall aims of the PIP are to enable:

- The legal and secure use of patient data for secondary purposes by the NHS (and other organisations involved in the commissioning and provision of NHS-commissioned care)
- NHS business to no longer use identifiable data in its non-direct care related work wherever possible
- NHS business processes to continue to be effective in supporting the day to day operation of the NHS.

The target of the project is the completion of this transition by April 2011.

Initial work of the Project included a questionnaire sent to all affected organisations. This provided valuable information on flows of patient data for secondary use purposes and identified practical issues to be addressed in implementation.

NHS organisations were required through the 2009/10 Operating Framework Informatics Planning Guidance to submit plans for implementation of pseudonymisation. A planning template and guidance (Ref 6), were provided to support development of high level plans by NHS organisations. The template and associated maturity model identified 14 key steps or requirements to be met as part of delivering local solutions and required completion and return to their relevant SHA and to PIP. The documents acted as a mechanism for generating high-level plans and raising organisational awareness of the need for de-identification of data used for secondary purposes.

© Crown Copyright 2009

In order to assess the starting point for each organisation and to be able to track progress, a Maturity Model was developed covering the 12 essential implementation steps. The planning template also gathered information on perceived issues and problems, which together with the baseline scores from the Maturity Model have been taken into account in preparing this guidance.

This paper provides the guidance referred to in the Planning Guidance on the implementation of data usage and governance for secondary uses and is provided in order that organisations can better understand the detailed requirements and for development of their detailed plans.

1.3 Related papers

This guidance builds on the earlier PIP documents

- PIP Implementation Planning Guidance
- PIP Maturity Model

The guidance is based upon a set of reference documents:

- Reference Document 1 – PIP Terminology, (Ref 1)
- Reference Document 2 – Business Processes and New Safe Havens, (Ref 2)
- Reference Document 3 – De-identification, (Ref 3)
- Reference Document 4 – Techniques, (Ref 4)

These reference documents have been written as standalone documents to provide coherent guidance in their own right in order to meet the needs of different audiences. It follows therefore that there is a limited duplication of content across the papers.

Much work has been already undertaken on information security across NHS and PIP has a stated aim to build on that work. Therefore, a separate paper on security has not been produced. Many aspects of security good practice are cited within the reference documents; and CFH publish Good Practice Guidance on relevant subjects, such as cryptograhic algorithms and password management.

1.4 Who should read this paper?

The PIP Implementation Planning Guidance indicated that PIP applies to:

- All commissioners - responsible for their own data usage activities and that the responsibilities of providers are reflected in contracts
- All providers - including Independent Sector Treatment Centres/Providers (ISTC/P) and third sector providers
- All shared services operating services that process patient level data for secondary use purposes
- Ambulance services
- Public Health Observatories (PHO) - for those uses of patient data not covered by Section 60/251 regulations for Public Health.
- SHAs – to ensure that appropriate data usage takes place in all organisations in the SHA's area (i.e. assurance, capacity, capability); to ensure conformance by any 'hosted services' for which they are responsible and may have identifiable patient data used for secondary purposes; to ensure conformance where the SHA may be a major stakeholder or participant in projects analysing data to support performance management and system reform.

© Crown Copyright 2009

The focus of the papers is on NHS commissioners and providers of services for NHS commissioned care. Where specific organisation type guidance for commissioners and providers appears in the guidance documents, then:

- PHOs should follow the guidance for commissioners;
- Ambulance services and ISTC/Ps should follow the guidance for providers
- Shared Services should follow the guidance relevant to the types of organisations to which they are providing services.

This paper and the others referenced above are aimed at the staff responsible for implementing de-identification and good practice in local data usage of patient data for non-direct care purposes.

These papers are therefore intended to provide information for a range of staff including:

- Informatics directors and managers with responsibility for implementing local data usage and governance
- Project managers with responsibility for managing the resulting implementation project
- IG staff for understanding the IG implications of de-identification of patient data when used for non-direct care purposes
- Managers responsible for business processes currently using patient identified data for non-direct care purposes and needing to make changes
- ICT management and staff for assessing the impact on systems and determining the means of and undertaking the implementation of the technical solutions.
- These papers are ***not*** intended for Chief Executives; the papers may be useful for Caldicott Guardians and Senior Information Risk Owners if they need to reference material in relation to secondary uses.

1.5 How to use this guidance

This guidance and its supporting documentation can be related directly to the essential implementation steps used in the Planning Guidance and Maturity Model as illustrated in **Figure 1**. Further information is given in Section 5 on accessing guidance for specific organisations.

Figure 1 Guidance and Implementation Steps

Implementation Step		**Main Guidance**	**Safe Haven & Business Process**	**De-identification**	**White Paper on Techniques**
R7	Data Management			√	
R8	Pseudonymisation Functionality			√	√
R9	New Safe Haven	√	√		
R10	Access Control	√	√	√	
R11	User Registration		√	√	
R12	End user applications			√	
R13	Business process change	√	√		
R14	Log & audit trails			√	

1.6 Patient label

Throughout this and the related papers, the term ‘patient label’ is used to describe the data item(s) that distinguishes one patient from another in a set of data; please note that this is purely for the purposes of clarity within the papers.

For identifiable data, the patient label may be the NHS Number (if present) but could be within an organisation, its Local Patient Identifier; another data item, or combination of data items, that can be used to uniquely identify one patient from another; in de-identified data sets, patient labels will vary but, for example, can be pseudonyms or table row numbers (the latter may not be unique as activity relating to the same patient may appear more than once in a table.)

1.7 Applying the Guidance

This document and its associated reference papers provide guidance on how to mitigate risk in order to balance risks associated with patient data with being able to effectively operate NHS business processes. The guidance is based on principles and is necessarily generic and cannot take into account the myriad of organisational arrangements affecting access to patient data, such as PAS facilities shared between multiple providers and commissioners. The application of the guidance on a local basis needs to be undertaken with due regard to the local circumstances.

In the event of uncertainty or requiring help, please check with the SUS Pseudonymisation websites[7] (including the FAQs) or contact the local SHA PIP Lead.

[7] **http://www.connectingforhealth.nhs.uk/systemsandservices/sus/delivery/pseudo** or **http://www.ic.nhs.uk/services/the-secondary-uses-service-sus/pseudonymisation-implementation-project**

© Crown Copyright 2009

2 Secondary Uses and Risk Mitigation

2.1 Secondary Uses

A high level definition of secondary uses is developed in the PIP Reference Paper 1 Guidance on Terminology (Ref 1). Secondary uses equates to non-healthcare medical purposes use within medical purposes as set out in *Confidentiality*: the NHS Code of Practice (Ref 5). In effect, secondary use of patient data is the use for purposes that do not directly contribute to the safe care of the individual concerned.

Purposes that directly contribute to the safe care of the patient are classified as primary uses and include care, diagnosis, referral and treatment processes together with relevant supporting administrative processes, such as clinical letters and patient administration, patient management on a ward, managing appointments for car; as well as the audit/assurance of the quality of the healthcare provided.

Confidentiality also clearly states that use of patient data for non-healthcare medical purposes must be 'effectively anonymised', that is in de-identified form unless it is with the patient's consent or otherwise covered in law, such as with approval under Section 251 of the 2006 NHS Act given by the National Information Governance Board (NIGB) Ethics and Confidentiality Committee (ECC).

It is necessary to distinguish between the two types of use in order to determine what data a user can see. Examples of secondary use of patient data are performance management, commissioning, contract monitoring; all of which do not require the identity of patients. There are functions within the NHS that use the same data sources for both secondary and primary uses. An example at PCT level would be for performance monitoring of Referral to Treatment (RTT) which should use de-identified data, but for organising care provision within 18 weeks, access to identifiable data is a primary use and therefore permissible by a suitably authorised member of staff.

Practical examples of secondary uses are given in the PIP Reference Paper 2 on Safe Haven & Business Processes Guidance (Ref 2).

2.2 Risk Mitigation and implementing de-identification

The aims of PIP are concerned with mitigating risks, in particular those associated with the use of identifiable data, and ensuring that patient data is managed in legal and secure ways. It is vital that clinical risk is not introduced or increased by introducing de-identification; for example failure in accurately identifying patients needing follow-up action after analysis of pseudonymised data, or inappropriately de-identifying data.

The risk of identification of patients, either inadvertent or malicious, when the data is used for secondary purposes, such as performing NHS business processes, must be minimised. This is because the identity of patients is not usually pertinent to secondary uses and if identity is pertinent then relevant channels to gain access to identifiable data should be pursued.

As indicated in *Confidentiality* it is necessary to achieve a suitable level of 'effective anonymisation' in the context of NHS business uses; this guidance sets out the steps necessary to reach such a level.

The aim is therefore to minimise risk to an acceptable level in the context of NHS business; there are a variety of techniques to achieve this whilst balancing risk reduction against the utility of the data and its purposes.

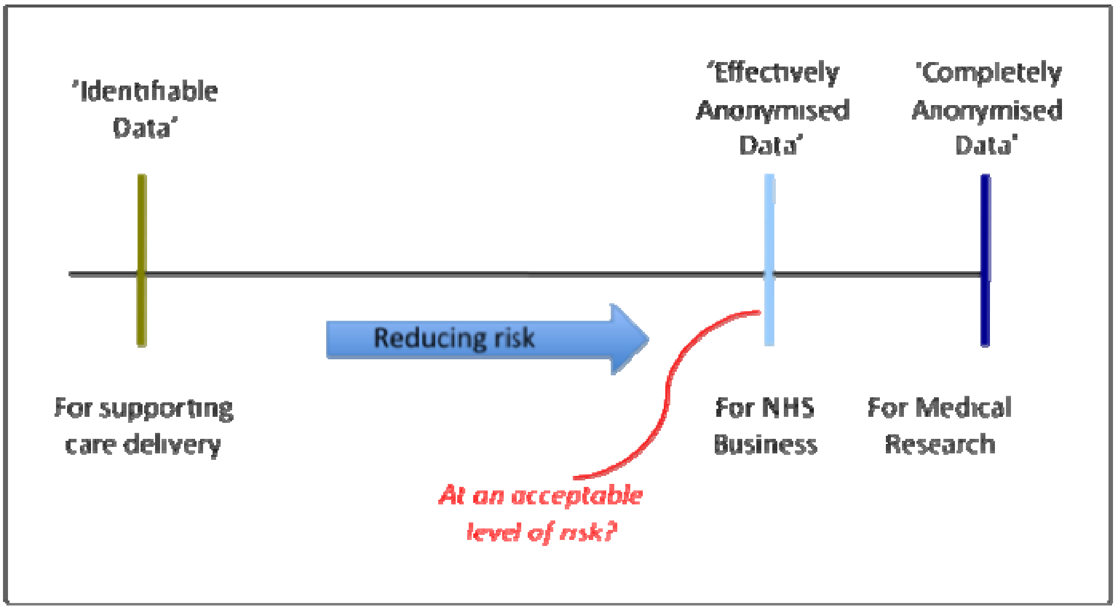
Achieving effective anonymisation requires actions to be taken on data handling and management regimes, modification of business processes and implementation of de-identification.

© Crown Copyright 2009

De-identification is the general term for any process of removing or reducing the association between a set of identifying data and the data subject in order to protect personal confidentiality. This can range from partial de-identification (where only some of the person identifiers such as NHS number, date of birth, postcode are removed) to complete (where all person identifiers are removed) and a range in between.

It is important that the steps taken to mitigate the risks associated with the use of patient data for non-direct care purposes are recorded, whether they are for the full implementation of de-identification as set out in the PIP documents or not. This is because of the changing expectation on the use of identifiable data in this way and the potential for external review through the regulatory mechanisms.

Figure 2 Risk Mitigation


'Identifiable Data'
'Effectively Anonymised Data'
'Completely Anonymised Data'
Reducing risk
For supporting care delivery
For NHS Business
For Medical Research
At an acceptable level of risk?

2.3 Local Implementation

Local organisations must decide on their approach to implementation to achieve the de-identification of person level data that is used for non-direct care purposes. These decisions will necessarily be taken in the context of competing and financial priorities, and reflect the wider risk management responsibilities of these organisations.

Each organisation that makes non-direct care use of patient data organises its business processes, departments and systems to meet its own needs. Therefore there is great variety in the local context for the non-direct care use of patient data. The work undertaken by the PIP Team has included attempting to gain an understanding of that variety in order to develop generic solutions and guidance to enable de-identification and supporting measures to be implemented.

There are different ways to implement the steps necessary to enable the non-direct care use of patient data to be undertaken with de-identified data. The responsibility for achieving this as part of an organisation's Information Governance obligations rests with the management of the organisation. The decisions on how to implement local data usage and governance using de-identification are for each organisation to take.

Please note that implementation of all processes for de-identifying, pseudonymising and re-identifying patients must be subject to adequate testing prior to live deployment in business processes.

© Crown Copyright 2009

3 Context for implementing Local Data Usage and Governance

3.1 Information Governance Context

The Caldicott Principles as set out in Table 1 apply to uses of patient data, including those for secondary purposes and these principles form a backdrop against which implementation of local data usage and governance takes place.

Table 1 Caldicott Principles

Principle 1: Justify the purpose(s)	Every proposed use or transfer of patient-identifiable information within or from an organisation should be clearly defined and scrutinised, with continuing uses regularly reviewed by an appropriate guardian.
Principle 2: Do not use patient-identifiable information unless it is absolutely necessary	Patient-identifiable information items should not be used unless there is no alternative.
Principle 3: Use the minimum necessary patient-identifiable information	Where use of patient-identifiable information is considered to be essential, each individual item of information in a data set should be justified with the aim of reducing identifiability.
Principle 4: Access to patient-identifiable information should be on a strict need to know basis	Only those individuals who need access to patient-identifiable information should have access to it, and they should only have access to the information items that they need to see.
Principle 5: Everyone should be aware of their responsibilities	Action should be taken to ensure that those handling patient-identifiable information, clinical and non-clinical staff, are aware of their responsibilities and obligations to respect patient confidentiality.
Principle 6: Understand and comply with the law	Every use of patient-identifiable information must be lawful. Someone in each organisation should be responsible for ensuring that the organisation complies with legal requirements.

3.2 NHS Business Needs

NHS business use of patient data depends on:

- Data being of sufficient quality to support the accurate business operations
- The ability to link records across provider organisations and across time periods
- Generating derivations.

The above means that some vital processing steps in the use of patient data have to be feasible; in particular processing to:

- Ensure adequate data quality for the business use to be effective

© Crown Copyright 2009

- Enable new important data to be derived from data items in the original records, such as area based deprivation indices, which are derived from full postcodes
- Ensure that the linkage of records relating to an individual patient is feasible and accurate. Individual patient activity records can be linked in relation to a patient's health issues in several ways to:
  - build spells from episodes of different types of activity (e.g. outpatient and admitted patient care) within a provider
  - build patient pathways from episodes of different types of activity across several providers
  - create longitudinal records within and between providers from episodes, spells, contacts and other events
  - create longitudinal records within a commissioner across providers from episodes, spells and pathways.

3.3 Secure environment

Local organisations must implement de-identification processes and solutions within a secure environment requiring compliance with the conditions set out below.

These processes and solutions will be required when the:

- Secondary purposes for which the data are being used are not covered by patient consent, Section 251 or other legal justification
- Data available to the organisation relates to patients or residents for which the organisation has responsibility and a duty of care or is acting as a data processor to the same ends

The organisations implementing de-identification processes and solutions must already have implemented:

- The mechanisms and facilities set out in the guidance within the PIP planning template, which requires the implementation and operation of local 'new safe havens' supported by rigorous access controls, and systems, which are only accessible to suitably registered users
- A data management regime that enables the storage of identifiable data with access for authorised users and protection from unauthorised end users.

The services in these organisations will be operated by staff who are bound by contractual obligations concerning patient confidentiality and related disciplinary regimes

These conditions build on and complement the existing information governance and security infrastructure and good practice that has been developed within the NHS. NHS organisations should look to ISO 27001/2 on Information Security[8] and the NHS CFH Good Practice Guidance[9] (GPG) for specific guidance on security aspects of their systems.

3.4 Outside the Secure environment

The PIP guidance has been developed on achieving effective anonymisation for supporting NHS business processes within the secure environment conditions set out above. Additional processing is required to reduce the levels of identifiability for secondary uses of patient data

[8] 8 ISO 27799 Health Informatics – Information security management in health using ISO/IEC 27002

[9] **9 http://nww.connectingforhealth.nhs.uk/infrasec/gpg**

© Crown Copyright 2009

outside the secure environment. These include further de-identification steps and rules on display of data.

3.5 Organisations covered

The requirement to implement de-identification processes and solutions applies to all members of the NHS family of organisations, which make secondary use of patient level data, (i.e. NHS commissioned care service providers including ambulance trusts and independent sector treatment providers, commissioners, shared services and public health observatories). It also applies to services provided by a 'data processor' on behalf of an organisation. These services should comply with the security and data management processes outlined above.

The requirement applies to all local configurations of systems, for example where:

- Organisations have implemented multiple systems from different vendors which support different business purposes; and,
- Data storage, processing and display takes place across multiple systems or subsystems, including in-house and commercially available products.]

Suppliers providing services through the Framework for procuring Support for Commissioners (FESC) should be considered as part of the commissioning organisation and are required to operate within the same guidance as the commissioner.

Other third party service suppliers, which are used by NHS Organisations to process and analyse data under a contract, which leads to the third party being classified as a 'Data Processor' under the Data Protection Act are expected to have implemented a trusted and secure environment. Contracts with such third parties should include a requirement party to conform to NHS guidance on pseudonymisation.

Similarly, where services are provided to PCTs by Public Health Observatories (PHO) and Quality Observatories (QO), these should have implemented a trusted and secure environment and conform to this guidance.

© Crown Copyright 2009

4 Implementation Guidance Framework

4.1 Implementation Guidance development

Different approaches for implementing de-identification of records for use for non-direct care purposes have been developed following extensive dialogue with staff using data or implementing and supporting systems and processes in NHS organisations, as well as with those responsible for developing NHS policy, including Information Governance. The approaches seek to balance the requirement to meet legal and security obligations with enabling the NHS to support the provision of services and undertake the business processes necessary to achieve the policy aims of the NHS, such as payment by results and referral to treatment targets.

The implementation guidance has been developed to meet the need for effective anonymisation and whilst supporting NHS business and to enable completion of implementation by March 2011.

4.2 Implementation Guidance rationale

The guidance is built on the following propositions:

- Data items cannot in themselves be classified as primary or secondary data, but are associated with direct and non-direct care purposes when they are used.
- The quality of data within healthcare records is of direct importance to the safety of patients. The results of any data quality analysis that is undertaken, whether it is associated with a primary or secondary use of data, are valuable and should be reflected in the source of the data, usually a system supporting direct care activities.
- There are legitimate reasons for clinicians to access identifiable data following processing of data for non-direct care purposes.
- There are multiple flows of data between NHS commissioners and providers of NHS commissioned care from which the data are used for non-direct care purposes.
- All flows between organisations are now required to be in encrypted form to minimise risks of inappropriate disclosure.
- Problems with data quality may be detected when records are being linked or data are combined or processed and resolution is needed; resolution is only feasible through reference back to the sources of the data. This requires some form of identification of the record or the subject of the record. Wherever possible this should be achieved preferably by referencing a record identifier; if this is not feasible, then an appropriate patient label, such as Local Patient Identifier or NHS Number, should be used.
- The capability does not currently exist to provide a comprehensive national system of pseudonymisation to support the variety and volume of flows of data. De-identification at the source systems is not feasible on a pan NHS basis. Neither is such a mechanism feasible given the current levels of data quality.
- Data flows between providers of care and commissioners therefore need to be in identifiable form because of their potential use for direct care purposes and for resolution of data quality issues. However, the flows must have the minimum possible identifiable data, such as NHS Number, date of birth and postcode, and exclude names and addresses.
- NHS IG policy requires that data flows of sensitive information between organisations must be encrypted for security purposes and to minimise risks to confidentiality.

© Crown Copyright 2009

- Non-healthcare medical business processes undertaken in the NHS must be carried out using de-identified data.
- Records must be de-identified before the data are used for secondary use purposes; which means that de-identification should take place on receipt of data and ahead of usage. This should take place within local NHS organisations.
- Records with identifiable data must be received, stored and managed in a controlled manner with the transition processes of data quality, linkage, derivations and de-identification being undertaken with minimal access to identifiable data.
- The transition processes to create de-identified data from identifiable data can themselves be construed as having non-direct care purposes.
- It is necessary therefore to seek permission under Section 251 of the 2006 NHS Act in order to ensure that these steps (outline above) have a clear legal basis. This is being pursued as part of the work of PIP.
- The project will build on the existing NHS IG infrastructure and standards and the CFH Good Practice Guidance.
- Overall the concern is with ensuring the security of data within, and across, organisations using patient data and on its transfer to external organisations. Guidance is therefore needed on how data can be transferred without a person or patient label (or other identifiers which could identify a patient with or without other data) to allow non-direct care business activities to be undertaken.

To support the implementation, PIP is developing a Standards Framework for De-identification of NHS Secondary Uses Data for the Information Standards Board for Health and Social Care and a DSCN should be issued early in 2010.

4.3 Implementation Guidance – out of scope

To clarify what is and is not included in the project's planned implementation, it is important to state what PIP is not attempting to implement. The reason for this is to dispel myths, genuine misunderstandings and possible items of disinformation that have developed during the project. As PIP has explored the development of suitable solutions, there have been many discussions and pursuit of differing routes, some of which may have led to potential confusion and mixed messages.

To confirm that:

- This guidance covers the de-identification of patient data, which is used for non-direct care purposes and is applicable to all commissioners and providers of NHS care including Foundation Trusts
- the requirements will be integrated into the next iteration of the IG Toolkit
- patient level data will continue to be transferred between NHS organisations (or those providing services via FESC or appropriate contractual arrangements) in identifiable and encrypted form
- PIP is concerned with the use of patient data within organisations in communications between organisations as components of operating NHS business processes.

Conversely:

- PIP does **not** require data submitted to SUS to be pseudonymised
- PIP does **not** require data transferred between NHS organisations (or those providing services via FESC or appropriate contractual arrangements) to be pseudonymised

© Crown Copyright 2009

- A single national set of pseudonyms will **not** be introduced other than those available in SUS reports and extracts.
- Data extracted by commissioners and providers from SUS does **not** all need to be pseudonymised.

4.4 Relating Implementation Steps to Guidance

The reference papers provide detailed guidance for implementing the changes necessary to move to undertaking secondary uses with de-identified data. Outlines of their contents are given below.

Reference paper 1 – Terminology.

This paper provides the legal and policy framework for secondary uses and the requirement for the use of 'effectively anonymised' data in secondary uses. The term Secondary Uses is defined and contrasted with primary uses and gives a high level view of the sets of techniques that can be used to enable patient data to be regarded as 'effective anonymised'.

Reference paper 2 – Business Processes and New Safe Havens

This paper builds on the definition of secondary uses from Reference Paper 1 and the need to use to de-identified data by considering the implications for NHS business purposes. A requirement arises for being able to identify patients as legitimate output from secondary use analysis and for accessing identifiable data as the only means of meeting data quality, record linkage and derivation needs. This leads to the introduction of the New Safe Haven as the means of controlling access to identifiable data for the identified needs.

The implications of the use of de-identified data and the New Safe Haven arrangements are then worked through to lead to guidance on business processes in general. Guidance is also provided for some specific business processes in response to issues raised by NHS organisations. This guidance will become web based and updated as developments occur and learning can be shared.

Reference paper 3 – De-identification

This paper builds on the material from Reference Paper 1 concerning methods of de-identification and relates this to the secure environment requirement in this guidance paper. The data items that act as identifiers are listed, together with the means of de-identifying such data and how and when data can be displayed.

Specific techniques are described for creating pseudonyms and for accessing identifiable data. The foregoing are brought together in a set of rules, supplemented by the requirements for logging and audit of access to identifiable data. Guidance is provided on specific issues, such as shared services as well as provider stand-alone and legacy systems. Again, this guidance will become web based and updated as developments occur and learning can be shared.

4.5 Specific Issues arising from the Guidance

There are some significant and specific requirements arising from the guidance as set out below:

- **PCTs** – separation of commissioner and provider data usage. De-identification and display of data for secondary use purposes is dependent on the context of use and accessibility of systems to identify patients. There are significant differences between providers and commissioners in the data allowed to be displayed – see Section 3.2 Reference Paper 3 De-identification. For this reason, there is a requirement to provide

© Crown Copyright 2009

different data displays for users in PCT provider arms from those for PCT commissioner arms.

- **Invoicing** - identifiable data may ***only*** flow from providers to commissioners' New Safe Havens. This means that invoice flows from provider finance departments to commissioner finance departments must not contain NHS Numbers, dates of birth and postcodes, as these are not relevant to financial aspects. Issues concerning the correct attribution of patients to practices and PCTs are matters of data quality and must be dealt with through the New Safe Haven process. This applies to regular transactions and non-contract activity. See Section 5.5 Reference Paper 2 Business Processes and New Safe Haven on inter-organisational communications.
- **Legacy systems** – there may be systems producing output used for secondary purposes, which cannot be modified to produce de-identified data. Possible ways of tackling this are provided in Section 5.3 Reference Paper 3 De-identification. However, if solutions are not feasible or not cost-effective, then such exceptions must be recorded and an exit strategy proposed.
- **Data Processors** – if an NHS organisation makes use of third parties to process data on their behalf, then it is incumbent on the NHS organisation to be clear whether the third party is a Data Processor as defined in the DPA or not. If the third party is a such Data Processor it must be registered as part of the NHS organisation's DPA registration; further the third party must conform to this guidance on the same basis and the NHS organisation itself. If the third party is not registered as a Data Processor, then identifiable data must not be supplied to it.

© Crown Copyright 2009

5 How to proceed

5.1 Overview

It is the responsibility of individual NHS organisations for implementing the measures to enable secondary uses of patient data in the NHS to be put on a safe legal footing. The guidance in this paper and supporting documents will enable the goal to be reached.

Each organisation will need to decide for itself how to undertake the implementation in relation to its own particular circumstances. The local decisions will relate to consideration of the following aspects amongst others:

- Type of organisation
- The starting point in terms of IG arrangements and facilities
- How local business processes operate
- Whether commercial systems are utilised
- Whether multiple systems are involved
- Whether a shared service is used
- Whether third party data processing is undertaken as Data Processors under the Data Protection Act.
- Whether in-house capacity and capability is available whether data is extracted and provided to third party organisations outside the secure environment and not acting as a Data Processor.

These issues are considered further below through looking at the actions required to implement effective anonymisation of patient data to support secondary use purposes. The actions set out in the next two sections reflect elements of the guidance set out previously in this document and in Reference Papers 2 and 3. These actions have been brought together to provide a coherent path through the elements of the guidance.

5.2 Strategic Issues

There are some strategic issues that impact on and potentially determine the specific implementation actions that need to be undertaken within an organisation and/or shared service. These are centred on the following:

- Involve Caldicott Guardians from the outset and in the finalisation of use of identifiable data and for authorisation of any New Safe Haven and its operation
- If the organisation is part of a shared service, determine how New Safe Havens and de-identification will operate, that is on a collective and/or single basis for constituent organisations, for example implementing de-identification at the shared service implies that New Safe Haven services/facilities may need to be based there too in order to undertake DQ and other transition activities
- If Data Processor services are used, determine with them how New Safe Havens and de-identification will operate
- Set up review to ensure secure environment elements can be set up for secondary use of patient data and determine actions to ensure local secure environment fully implemented.
- If there are external data flows outside the secure environment, then facilities for generating multiple pseudonyms are required

© Crown Copyright 2009

- (mainly relevant to providers) Review to check if there are any legacy systems for which implementing pseudonymisation or data extraction for pseudonymisation elsewhere are not feasible/cost effective; document these as exceptions
- (providers) Review to see if there are stand-alone clinical systems that will continue to operate in stand-alone mode
- Establish protocols for communication between organisations with routine high volume data flows on identifiable and de-identified data in order to ensure safe haven to safe haven communications
- Determine how to review business processes
- Review invoicing processes
- Determine approach to de-identification – in-house, through existing secondary use systems, supplier, commercial third-party supplier.

5.3 Implementation Actions

A set of actions will be generated from consideration of the strategic issues as above. These actions will overlap with the actions set out below generated from within or implied by this paper and the reference papers.

Please note:

- The specific actions required by an organisation are dependent on local circumstances and not all of these may apply to each organisation.
- The list below is not exhaustive, but covers many of the major actions; it is also not implying these actions should be undertaken in the sequence listed or by all organisations. Some actions may be specific only to Providers or to Commissioners
- The list is provided as checklist to help organisations work through potential actions.

Organisational issues

- Organisational awareness for data users of changes to the local data usage and governance, impacting on access to identifiable data and business processes.
- Involve the Caldicott Guardian in planning and implementing changes.

Secure environment

- Check that security facilities conform to secure environment standards.
- Ensure access controls can cope with differentiating roles and the separation of access to de-identified and identified data.

Review inter-organisational communications

- Identify regular high volume inter-organisational communications and determine need for protocol to govern data exchanges.
- Identify safe havens with which communications are expected to and from.
- Invoicing – providers – ensure that invoices do not contain patient identifiers, namely NHS Number, Date of Birth and postcode, and that invoices are for activity unless it is safe to use LPI as a fully quarantined data item.
- Invoicing – commissioner – ensure that invoice receipt and payment processes are modified to operate without patient identifiers.

© Crown Copyright 2009

Review business processes

- Separate out business processes involving data quality checking and determine how these will be managed through the New Safe Haven mechanism – including invoicing, OATS and high cast cases,
- Consider WCC business processes currently undertaken and plan for any future known/expected developments.
- List and review business processes using Type A, B and C model – e.g. RTT & waiting list management.
- Identify and modify business processes where identifiable data no longer can be supplied, i.e. Type B and Type B parts of Type C business processes.
- Determine the approach to spatial analysis.

New Safe haven set up

- Identify points of receipt of identifiable data for systems where secondary use is expected to be made.
- Identify and reorganise functions – covering DQ, derivations, linkage and de-identification.
- Define extent of DQ operations required in order to define New Safe Haven.
- define in terms of functions, posts/people, facilities
- modify systems and storage to support New Safe Haven functions and operations
- Identify & register staff authorised to utilise identifiable data
- Implement New Safe Haven security, including access controls
- Authorise staff via Caldicott Guardian and Registration Authority processes
- Re-org system ops and access controls to support dq, derivations, linkage and
- Sign-off by Caldicott Guardian of arrangements for access to identifiable data

De-identification Implementation

- Stage 1 – strategy/design
  - Define environment and data sources (e.g. shared PAS) in order to define de-identification regime for identifiable data items (e.g. LPI)
  - Determine means of de-identification (as per Section 2.4 of Reference Paper 2)
  - Determine method of pseudonymisation if required
  - Determine means of enabling access to identifiable data for authorised users
  - Determine means of logging and auditing of access to identifiable data
  - Determine approach to missing NHS numbers
  - Design secure storage for pseudonymisation parameters/keys
  - Design secure storage and access controls for look-up tables (if used)
  - Design secure storage and access controls for holding identifiable data logically separate
  - Define protocol for local display of identifiable data items (as per Section 3.2 of Reference Paper 2) to cover different output types, such as reports, extracts, user configurable 'slice & dice' cubes.

© Crown Copyright 2009

        - Note 1 – the display of data items is related to the users' access to other sources of patient data, e.g. do not use LPI at a commissioner if there is shared access to the relevant PAS.
        - Note 2 – if relevant provide 'health warnings' concerning use of specific data items, such as ethnic category.
    - Define protocol for provision of data outside the secure environment (if relevant)
- Stage 2 - implementation
    - Implement facilities, system changes etc arising from Stage 1
    - Cease extracts of identifiable data for end users

Users

- Register users of patient data by use type – that is access to pseudonymised or authorised to access identifiable data or both in relation to specific roles
- Determine approach on dual/multiple roles in general and in particular in small teams
- Modify access rights for users – default is access to pseudo data unless authorised to view identifiable

Specific Issues

- Review legacy systems – identify, document exceptions, develop exit strategy
- Providers - Review standalone clinical systems
- Providers – Clinical teams
- Providers – secondary uses
- Shared PAS systems – across providers and commissioners

© Crown Copyright 2009