Warwickshire County Council - Pseudonymisation Procedure

Accountable Officer:	Tonino Ciuffini, Head of Information Assets
Author	Gareth Wrench, Public Health Intelligence Manager
Approved by:	Information Governance Steering Group
Date approved: April 2015	
Issue date: 12 March 2015	
Review date: April 2017	
Person responsible for instigation:	Andrew Morrall, Corporate Information Manager
Related Legislation:	Data Protection Act 1998 Common Law Duties of Confidentiality
References:	NHS Code of Practice on Confidentiality NHS Connecting for Health IG Toolkit NHS Operating Framework 2010/2011 Caldicott Committee Report 1997
Related Policies:	WCC Information Strategy WCC Information Governance Management Framework Information Safe Haven and Handling Procedure

Approvals

V1	Corporate Information Manager	March 2015
V1	Information Governance Steering Group	April 2015

Warwickshire County Council - Pseudonymisation Procedure

Contents

Warwickshire County Council - Pseudonymisation Procedure

1. Introduction

Pseudonymisation and de-identification is concerned with enabling Warwickshire County Council to undertake secondary use of personal confidential data (PCD) in a legal, safe and secure manner.

The overall aim of this procedure is to facilitate:

- The legal and secure use of PCD for secondary purposes by the County Council (and other organisations involved in the commissioning and provision of health and social care services).
- System-wide integration to enable relevant information to be shared between health and social care partner organisations.

These procedures are **mandatory** and must be followed by all staff as part of the council's Information Governance Framework the standard for managing information in the council and is one of the linked procedures in the Information Compliance policy aimed at all staff

This procedure applies to all staff including: council employees, contracted third party organisations and individuals (including agency staff), students/trainees, secondees, and staff or partner organisations with approved access (hereafter referred to as staff) who use PCD for non-healthcare medical purposes (secondary uses).

2. Scope

This procedure is concerned with the security of PCD when used for non-healthcare medical purposes (secondary uses) whilst balancing the needs of the County Council to perform its everyday business functions.

3. Definitions

Pseudonymisation

The technical process of replacing person identifiers in a dataset with other values (pseudonyms) available to the data user, from which the identities of individuals cannot be intrinsically inferred. For example, replacing an NHS number with another random number, replacing a name with a code or replacing an address with a location code. Pseudonymisation may be reversible or irreversible and forms a core element of Secondary Use Services. This is illustrated in the examples below:

Pseudonymisation – one-off	Use encryption or randomisation technique to generate unique random number to replace a NHS Number.

Pseudonymisation – repeated & consistent	When pseudonymisation techniques are consistently applied, the same pseudonym is provided for individual patients across different data sets and over time. This allows the linking of data sets and other information which is not available if the PCD is removed completely.

Anonymisation / De-Identification

Staff must only have access to data that is necessary for the completion of the business activity which they are involved in. This is reflected in Caldicott Principles; *access should be on a need to know basis*. This principle applies to the use of PCD for secondary or non-direct care purposes. By de-identification, users are able to make use of patient/client level data for a range of secondary purposes without having to access the identifiable data items.

The aim of de-identification is to obscure the identifiable data items within the person's records sufficiently that the risk of potential identification of the subject or a person's record is minimised to acceptable levels, this will provide effective anonymisation. Although the risk of identification cannot be fully removed this can be minimised with the use of multiple pseudonyms.

De-identified data should still be used within a secure environment with staff access on a need to know basis.

De-identification can be achieved by:

- Removing patient identifiers;
- The use of identifier ranges, for example; value ranges instead of age;
- By using a pseudonym.

Personal Confidential Data (PCD)

Personal Confidential Data (PCD) is any information that can identify one person. This could be one piece of data, for example, a person's name or a collection of information including name, address and date of birth.

Strong, direct examples of PCD include:

- Name
- Initials
- Address
- Postcode
- Date of birth
- Date of death

- NHS Number
- Local system identifiers
- National Insurance Number

Weaker, indirect examples of PCD may include:

- Ethnicity
- Rare diagnosis

Primary Use of Data

Primary use of data is when information is used for healthcare medical purposes. This would directly contribute to the treatment, diagnosis or the safe care of the individual. This also includes supporting administrative processes, the clinical audit/assurance of the quality of healthcare service provided, drug safety and public health observation.

Secondary Use of Data

Secondary use of data is for non-healthcare medical purposes which do not directly contribute to the safe care of the individual. This includes performance management, commissioning, research purposes, contract monitoring and reporting facilities, all of which do not require the identity of the patient or service user. When PCD is employed for secondary use, this should be limited and de-identified so that the secondary uses process is confidential. There are some exceptions to this, for example, Court Reports and Police Reports.

4. Principles

a) It is NHS policy and a legal requirement that when PCD is used for purposes not involving the direct care of the patient, i.e. secondary use, the patient should not be identified unless other legal means hold, such as the person's consent or Section 251 approval. This is set out clearly in the NHS policy and good practice guidance document 'Confidentiality: the NHS Code of Practice', which states the need to 'effectively anonymise' data prior to the non-direct care usage being made of the data.
b) Staff must only have access to the data that is necessary for the completion of the business activity which they are involved in.
c) This procedure is in line with the NHS Code of Practice on Confidentiality, the Information Governance Toolkit version 11 and the NHS Operating Framework 2010/2011.

5. When should pseudonymisation be used?

When PCD has agreed to be shared with partner organisations for secondary use purposes, it should be suitably pseudonymised using appropriately robust methods so that it is not possible to identify individuals. Effective controls should be in place to prevent any potential re-identification.

To effectively pseudonymise data the following actions must be taken:

- Each field of PCD must have a unique pseudonym;
- Pseudonyms to be used in place of NHS Numbers and other fields must be of the same length and formatted on output to ensure readability. For example, in order to replace NHS Numbers in existing report formats, then the output pseudonym should generally be of the same field length, but not of the same characters; i.e. 5L7 TWX 619Z. Letters should be used within the pseudonym for an NHS number to avoid confusion with original NHS numbers;
- Where used, pseudonyms for external use must be generated to give different pseudonym values in order that internal pseudonyms are not compromised.
- The secondary use output must, where pseudonyms used, only display the pseudonymised data items that are required. This is in accordance with the Caldicott Guidelines.
- Pseudonymised data should have the same security as PCD.

6. Safe Havens

Safe Haven is a standard term in use throughout health and social care, the NHS and its partners. It is a standard laid down by the Health and Social Care Information Centre as part of their duty under the Health and Social Care Act.

An Information Safe Haven is the outcome of a set of procedures that ensure the safety and secure physical and electronic handling of personal data whilst the information is located within the council and covers flows of such information within it and to other external partner agencies.

An organisation's Safe Haven procedure provides the guidance regarding the security of transferring information via fax, post, telephone and email. The organisation's Safe Haven procedure should incorporate the new Safe Haven principles.

The new Safe Haven principles include the concept of restricting access to identifiable data which is required to support the pseudonymisation process of de-identifying records. The new Safe Haven applies to the security of patient/client information systems and databases.

Patient/client information systems and databases must be held within an electronic safe haven whereby access is limited and password controlled for each authorised user.

Access to a new Safe Haven will be given by the WCC Information Assets department once a completed authorisation form has been received from the relevant Information Asset Owner.

A list of authorised users will be maintained for each Safe Haven database/system by the appropriate Information Asset Owner and a full list maintained by WCC Information Assets. This list will be regularly reviewed by the Information Asset Owners as part of the Senior Information Risk Officer reporting process.

7. Statement/Objective

It is a legal requirement that when PCD is used for purposes not involving the direct care of an individual, i.e. ‘Secondary Uses’, the individual should not be identified unless other legal means hold, such as the person's consent or Section 251 approval. This is set out clearly in the NHS policy and good practice guidance document 'Confidentiality: the NHS Code of Practice', which states the need to 'effectively anonymise' data prior to the non-direct care usage being made of the data.

Data cannot be labelled as primary or secondary use data - it is the purpose of the disclosure and the usage of the data that is either primary or secondary.

This means that even where it is justifiable to hold data in identifiable form, it becomes essential to ensure that only authorised users are able to have identifiable data disclosed to them.

The use of the NHS Number is Department of Health policy within Adult Social Care as well as in the NHS; its use is being supported by the NHS Number Programme. For services and organisations that operate across health and social care, the NHS Number is the key identifier of individuals and with greater access to NHS sourced data becoming available to social care organisations. The use of personal data in social care is governed by the same legal and policy frameworks that guide the use in the NHS and the Strategic Framework for De-identification covers social care as well as NHS. This means that secondary use of personal information within social care should operate against the principles and methods set out in this guidance.

8. Roles & Responsibilities

All databases, systems and spreadsheets holding PCD should be recorded on the Information Asset Register held by the Information Governance Team. Information Asset Owners (IAOs) have the responsibility of ensuring that the data extracted from the assets for which they are responsible are pseudonymised, or de-identified, when data is provided for secondary use.

Caldicott Guardians take the lead in patient/client confidentiality issues supported by the Corporate Information Manager.

The Corporate Information Manager is responsible for ensuring adherence to the Safe Haven Procedure. Any issues with de-identification of data, or breaches of data security should be reported by IAOs to the Information Governance Steering Group via the Corporate Information Manager.

Owners of Safe Havens will be responsible for ensuring that only staff with a genuine business need have access to personal identifiable information.

Each team must ensure that all data is handled in line with this procedure. When ad hoc requests for information are received the team member dealing with the request must establish what the information will be used for in order to categorise the use as either primary or secondary.

The Safe Haven comprises the facilities to restrict access by authorised users to personal identifiable data for the purpose of supporting de-identification, which in turn means that:

- The facilities can only be used by authorised staff sufficient to perform the functions and provide cover and back-up to ensure continuity of service.
- The systems (or sub-systems) used for the data transition processes must have appropriate access control mechanisms to restrict access to authorised users for the specific purpose of supporting de-identification processes.

Safe havens must be hosted on secure, restricted access parts of the WCC network.

The above locations are restricted to a named set of users, access being controlled by Warwickshire County Council network ID and password authentication log on.

Managers will ensure that access to the Safe Haven is kept in line with staff movement and any changes to a person's role will be reflected in a change of access to data repositories.

9. Data Presentation

Care should be taken in any analysis or presentation of data, that individuals cannot be personally identified. Disclosure methods may need to be employed where very small numbers are present, e.g. numbers less than five have been suppressed.

10. Training and Awareness

Effective pseudonymisation processes depend upon robust information governance and effectively trained staff who understand the importance of data protection and confidentiality.

The annual Information Compliance training that is delivered to all staff will reference the principles of Pseudonymisation.

The Information Management team will raise awareness of this procedure.

Advice and support around this procedure will be provided by the Corporate Information Manager.

11. Monitoring, Auditing, Reviewing & Evaluation

Managers are responsible for ensuring their staff comply with this Procedure.

Information Asset Owners should review and take into account the distribution of data from the asset and ensure that this procedure is adhered to.

A review of this document will be conducted every two years or following a change to associated legislation or national / local terms and conditions of service.

Any possible breaches to this procedure or data loss will be reported in line with the WCC Incident Reporting procedure.