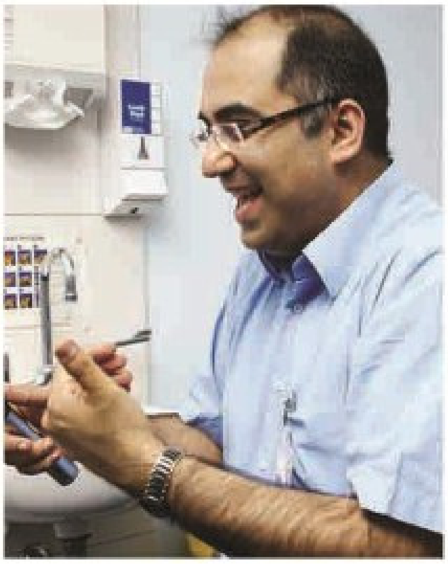
NHS England

Privacy Impact Assessment

Risk Stratification

IG Taskforce Consultation Paper CP-02

March 2014



THE NHS CONSTITUTION
the NHS belongs to us all

NHS England INFORMATION READER BOX		
Directorate		
Medical	Operations	**Patients and Information**
Nursing	Policy	Commissioning Development
Finance	Human Resources	

Publications Gateway Reference:	01335
Document Purpose	Guidance
Document Name	Privacy Impact Assessment - Risk Stratification
Author	W. Gowing
Publication Date	March 2014
Target Audience	CCG Clinical Leads, CCG Accountable Officers, CSU Managing Directors
Additional Circulation List	
Description	This privacy impact assessment is intended to support the work of NHS England in undertaking the risk stratification process for NHS patients.
Cross Reference	None
Superseded Docs (if applicable)	None
Action Required	N/A
Timing / Deadlines (if applicable)	**N/A**
Contact Details for further information	Stuart A Notholt Information Governance Communications Lead Phone 07796994375

Document Status

This is a controlled document. Whilst this document may be printed, the electronic version posted on the intranet is the controlled copy. Any printed copies of this document are not controlled. As a controlled document, this document should not be saved onto local or network drives but should always be accessed from the intranet

Gateway publication number: 01335

Privacy impact assessment – risk stratification

Considering and reducing the use of personal confidential data for the purposes of risk stratification

Information Governance Taskforce Consultation Paper CP-02

First published: March 2014

Prepared by Information Governance Taskforce

Contents

Consultation paper

This is a consultation paper. Readers are invited to comment on the areas covered in this document.

Please visit

http://www.england.nhs.uk/ourwork/tsd/ig/ig-consultations/

for details of how to contribute to this consultation.

Equality statement

Equality and diversity are at the heart of the NHS strategy. Due regard to eliminate discrimination, harassment and victimisation, to advance equality of opportunity, and to foster good relations between people who share a relevant protected characteristic (as cited in under the Equality Act 2010) and those who do not share it, has been given throughout the development of the policies and processes cited in this document.

1. Executive Summary

1.1. Introduction to privacy impact assessment and purpose of this paper

Privacy impact assessments (PIAs) are required by the Cabinet Office for information and communications technology projects in order to assess the risks to the privacy of individuals and the mitigation of such risks.

This paper sets out the benefits to arise from risk stratification, together with the risks in relation to the use of personal confidential data by health service organisations in risk stratification, and indicates how the identified risks will be mitigated and managed through a national approach and on a local basis.

This is also likely to be of interest to the independent sector risk stratification tool providers.

1.2 Introduction to risk stratification and short term issues

The purpose of risk stratification is to enable clinical commissioners to target specific patient groups and enable clinicians with the duty of care for individual patients to offer appropriate interventions.

Risk stratification is a process that can help determine which people in a population are at high risk of experiencing particular events, such as unplanned hospital admissions. As such, risk stratification is defined as a medical purpose, namely preventative medicine, supporting the provision of care and treatment, and the management of health and social care services.

Having provided advice on how risk stratification might be undertaken without using identifiable data, it has become apparent that in terms of the current available tools it is not feasible to conduct risk stratification without personal confidential data (PCD). Consequently NHS England sought and has obtained temporary support under the Section 251 regulations (see section 1.13.1) to set aside the common law duty of confidence, to enable personal confidential data to flow to the existing tool providers for this purpose.

Given the above context and the fact that risk stratification is a form of profiling, it is imperative for the public to be aware of the processing of their data for risk stratification purposes and their right to object to such processing, and further that they are aware of the risks in such processing and how those risks will be minimised.

1.3 Key findings from the PIA

The key findings are that

- There is a temporary legal basis for collecting and processing specific data for risk stratification; the statement and justification of the benefits having satisfied the Confidentiality Advisory Group (CAG) and the Secretary of State

- There are well defined and controlled processes for undertaking risk stratification
- A range of risks have been identified, together with suitable information governance controls to mitigate such risks
- There is a need to publish and actively disseminate privacy notices to inform the public of the use of their data for risk stratification and how they can object to such use
- Risk stratification will be undertaken on a local basis across the NHS in England and it is a compliance requirement of the Section 251 approval that a local privacy impact assessment must be undertaken.

1.4 Maintenance of the privacy impact assessment

NHS England aims to fulfil its statutory roles and functions efficiently and effectively, supporting commissioners in their work. Protection of privacy is fundamental to all that we do. This privacy impact assessment will be reviewed in step with the timing of the reviews undertaken by CAG in relation to the decision to approve setting aside the common law of confidence taken under Section 251 of the Health and Social Care Act 2006 – see Section 1.13.1. It is expected that, with experience of the operation of risk stratification, it may be possible to reduce the volumes and types of data required. Although it is also recognised that the technical structure of electronic health records may also constrain what is feasible in the migration to using more fully pseudonymised data.

2. The purpose of a privacy impact assessment

Privacy impact assessments were launched in the UK by the Information Commissioner in December 2007 and were mandated by the Cabinet Office for information and communications technology (ICT) projects following the Data Handling Review of June 2008[1].

A privacy impact assessment is a methodology to identify, assess, mitigate or avoid privacy risks. It describes the functions of the organisation to enable the reader to assess for themselves what may be considered a potential impact on their privacy, but it also goes on to explain what the organisation will do to protect individuals' privacy, and to identify solutions.

Risk stratification can use data about the health care services provided to individual patients by health care providers and general practice, either separately or in combination, to assess and predict future healthcare needs based on previous interactions. In order to enable future healthcare support to be provided at individual level, it is necessary to use personal confidential data at an appropriate point in the process.

There is a legal basis and a specific method through which risk stratification can be undertaken which maximises the protection of patient data. However, this is dependent upon the development of facilities, capabilities and capacity within the Health & Social

[1] http://www.ico.org.uk/about_us/consultations/~/media/documents/library/Corporate/Research_and_reports/pia-executive-summary.pdf

Care Information Centre (HSCIC), necessitating a short-term solution outside the legal framework provided for the HSCIC in the Health and Social Care Act 2012.

NHS England has, therefore, applied to the Secretary of State via the Confidentiality Advisory Group for approval of the "Disclosure of commissioning data sets and GP data for risk stratification purposes to data processors working on behalf of GPs" submitted for approval under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002 to process patient identifiable information without consent.

Approval has been made on a provisional basis, subject to compliance with specific conditions, for a six-month period.

The scope of this PIA will cover risk stratification and associated processes as covered by the Secretary of State's approval for the six-month period and any potential subsequent extension prior to the longer-term solution being implemented.

This privacy impact assessment:

- Describes the purpose and objectives of risk stratification
- Describes risk stratification processes and data management in the short term
- Assesses the potential implications for privacy
- Explains what NHS England will do to protect privacy
- Sets out what NHS England will require commissioners and data processors to do

3. What is Risk Stratification?

3.1 Risk Stratification Overview

The overall aim of the use of risk stratification[2] is to enable clinical commissioners to target specific patient groups and enable clinicians with the duty of care for the patient to offer appropriate interventions.

To enable this aim, risk stratification is a process that can help determine which people in a population are at higher than average risk of experiencing adverse events, such as unplanned hospital admissions, that are simultaneously: undesirable for patients; costly to the health service; and potential markers of low-quality care[3]. As such, risk stratification falls in legal terms under the following medical purposes - preventative medicine, and the provision of care and treatment, and the management of health and social care services - as defined within Section 251(12) of the NHS Act 2006.
Risk stratification is used to

[2] Risk stratification – a range of expressions may be used to describe risk stratification and related activities; these include risk profiling, risk prediction, risk modelling, predictive modelling and predictive risk modelling.
[3] See Information Governance and Risk Stratification: Advice and Options for CCGs and GPs.

- understand the characteristics of a local population (known as “risk profiling of a population” or “risk stratification for commissioning”)
- identify individual patients who are at risk of adverse outcomes such as unplanned hospital admissions, and who may benefit from additional preventive support such as that provided by community matrons (known as “risk stratification for case finding”).

To do this, the risk stratification process uses statistical analysis tools and models to analyse historic information such as age, gender, diagnoses, and patterns of hospital use to provide the basis of their predictions. Some models (e.g., PARR and PARR-30) use a combination of hospital data and geographical data such as the Index of Multiple Deprivation. Other models (e.g., the Combined Predictive Model) use primary care data derived from GP practice systems in addition to hospital data as the basis of their predictions.

A major use of risk stratification is to support long term condition management, as it has been found that better outcomes are achieved if the tools use applied statistical modelling based on 3-5 years of historical data to predict probability based on multiple risk factors, including hospital admissions and medication usage. For the tools to provide statistically robust risk score, the combined use of primary care and secondary care data in de-identified form is preferred.

The adoption of risk stratification by Clinical Commissioning Groups (CCGs) has been widespread, with variation in data specification and how risk factors are then used in the statistical calculation of risk. It is clear that the bespoke nature of risk stratification programmes is the result of much local clinical debate and tailored to support the priorities and designs of intervention programmes at a local level (e.g., reducing inequalities, prevention of re-admissions, long term conditions, frail elderly and proactive screening).

The two types of risk stratification, (i.e. one, commissioning and two, case finding) utilise the same source data and are undertaken within a single set of tools.

3.2 The Risk Stratification Process

Whilst the precise mechanism and processes used within in each CCG may vary, the two types of risk stratification operate on the same sources of data. The data sources and basis for disclosure of the PCD for the preliminary processing to combine and process primary care and secondary care data are:

1. commissioning data sets, which are disclosed from the HSCIC under s261(4) of HSCA 2012 Act
2. general practice (GP) data sets from GP systems under the instruction of GPs as data controllers.

Data are only disclosed to data processors if they meet the standards set out by HSCIC (as a minimum they will meet Stage 1 accredited safe haven (ASH) standards) or independent third parties acting as data processors where they fully meet IG Toolkit Level 2 and valid HSCIC data sharing contracts and HSCIC data sharing agreements are in place.

The data specifications are bespoke to each individual tool and have been agreed locally as part of the set up for risk stratification programme.

The data processors are either ‘in-house’ (55%) or commercial third party providers, which make up the rest with 11 main providers.

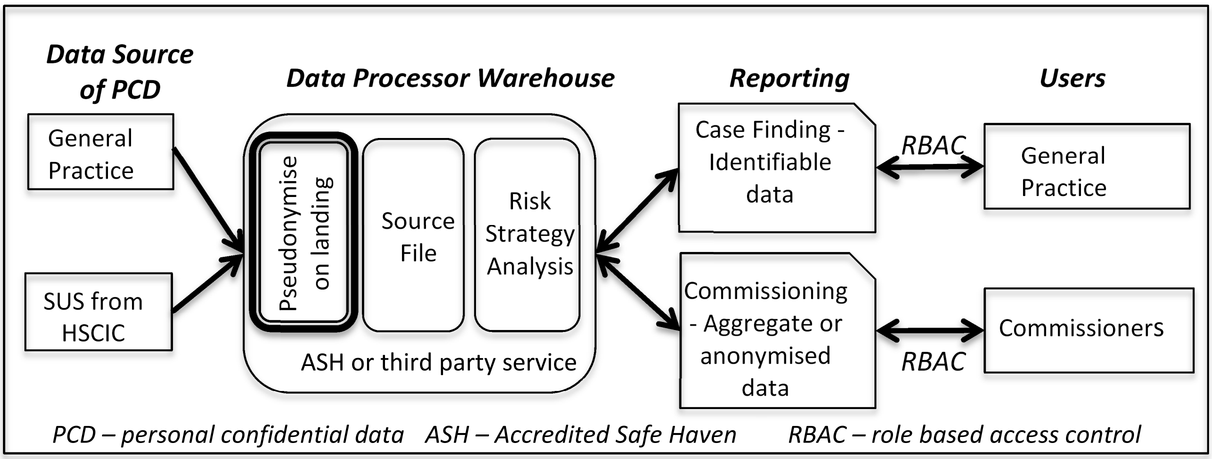
Patient records are pseudonymised on landing into a secure staging area, prior to feeding the pseudonymous data into the risk stratification tool. Some providers have opted to use a weakly[4] pseudonymised feed of the GP and secondary care data into the tool, as an alternative.

Almost all the organisations undertaking risk stratification have converged to using a closed system or ‘black box’ approach. They are all using role based access controls to separate the data processing and risk modelling as a closed system to allow clinicians with a direct relationship with patients to review and decide what intervention or pathway is most appropriate.

3.3 Risk stratification process overview

A generic model of the processes and data flows for the two uses of risk stratification is shown in Figure 1. There will be variations on this diagram depending on whether third party suppliers of risk stratification facilities are contracted to provide services or if the whole process is undertaken ‘in-house’. It should be noted that the same requirements in terms of security and access controls will apply as indicated above.

Figure 1 Generic Risk Stratification Process


Data Source of PCD
Data Processor Warehouse
Reporting
Users
General Practice
SUS from HSCIC
Pseudonymise on landing
Source File
Risk Strategy Analysis
ASH or third party service
Case Finding - Identifiable data
Commissioning - Aggregate or anonymised data
RBAC
RBAC
General Practice
Commissioners
PCD – personal confidential data ASH – Accredited Safe Haven RBAC – role based access control

4. Privacy issues of risk stratification

4.1 Risk stratification types

[4] Weakly pseudonymised data for risk stratification is defined as the following data elements NHS Number as the single identifier and include age, partial postcode, presence of date of death and sensitive items of gender and ethnicity

As indicated earlier, risk stratification can be split into two different purposes and two different types of output, namely:

- Risk stratification for commissioning
- Risk stratification for case finding.

Whilst there are different outputs, the results of the analysis are derived, as illustrated above, from the same sets of data arising from the flows of data from practices and from HSCIC for secondary care.

The flows, processing and use of patient level data can be split into five steps, namely

- Collection of data from general practice
- Collection of data from Secondary Uses Services within HSCIC
- Processing of data in Accredited Safe Havens (ASHs) or contracted third parties
- Provision of data to commissioners
- Provision of data to general practice.

The privacy implications of these five steps are covered in the following sections.

4.2 The collection/extraction of personal confidential data for risk stratification

4.2.1 Legal and constitutional basis

The context for the processing of patient level data is based on

1. The Health and Social Care Act 2012 providing a legal basis for the extraction of personal confidential data by the HSCIC in certain circumstances. The Act sets aside the requirement under the common law duty of confidence to seek patient consent to obtain the data[5].

2.Section 251 of the National Health Act 2006 also sets aside the requirement under the common law duty of confidence to seek patient consent to obtain the data subject to the approval of the Secretary of State.

3. Whilst common law duty of confidence may be set aside, the relevant requirements under the Data Protection Act 1998 continue to apply – in particular, the fair processing principle, which means that patients must be made aware of how their data are processed for risk stratification.

4. The Data Protection Act 1998 requirements. The first principle requires personal data to be processed fairly and lawfully and for one of the conditions in Schedule 2 to be met; and in the case of sensitive personal data, for one of the conditions in Schedule 3 also to be met. The purpose of risk stratification meets the following conditions:

[5] Under Sections 259 and 256 of the Health and Social Care Act 2012

a. Schedule 2 – the processing is necessary for the exercise of any other functions of a public nature exercised in the public interest by a person

b. Schedule 3 – the processing is necessary for preventative medicine and the management of healthcare services, and is undertaken by a person who in the circumstances owes a duty of confidentiality which is equivalent to that which would arise if that person were a health professional.

5. The NHS Constitution sets out a commitment giving patients the right to have their objections considered in relation to the disclosure of their information where the legal basis for disclosure is permitted in statute. Additionally, the Secretary of State has given a policy commitment that patients have the right to object to information about their identity in relation to personal confidential data leaving their GP practice and, other than in exceptional circumstances, to having their objections upheld in line with the commitment by the Secretary of State for Health in April 2013[6].

6. This commitment also applies to control of data about patients' identity leaving the HSCIC, although this can only apply to data collected after 1 January 2014.

The second step reflects NHS England's efforts to provide greater transparency and to enable patients to have greater control over the identifiable information held about them and is part of the care.data programme, to mitigate against the risks of loss of trust and loss of data.

The principle in point 3 above is important as it provides a mechanism for patients to protect their privacy and confidentiality in relation to primary and secondary care data if they so wish to do. This is because the data collection should occur within the healthcare provider and checks made against national systems (such as NHAIS and the Spine compliant systems) to confirm the relevant GP. The healthcare provider should ensure that any objections made by the patient are appropriately considered. The difficulties in doing so are recognised, in that currently provider systems do not usually have the means to record objections and certainly not in a systematic way that could be communicated effectively with other bodies.

In order to ensure that patients are aware of how data are processed for risk stratification as one of a set of uses of patient data and to ensure they are aware they can object, appropriate privacy notices measures will be undertaken as part of meeting Principle 1 of the Data Protection Act 1998 requirements on fair processing.

To support transparency and increase public awareness of the uses of their data and of their right to object, relevant privacy notices pertaining to risk stratification should be actively disseminated where data are collected, e.g. in practices, secondary care establishments and by the organisations undertaking the collection and usage, such as on the websites of practices and relevant CCGs.

[6] At the launch of the Caldicott Information Governance Review Report 26 April 2013

Organisations should develop their own communications strategy to consider how best to inform the public about the uses of their personal and confidential information and to support staff in giving this information and dealing with questions.

Whilst privacy notices reduce risks of processing data without the knowledge of patients, the extraction of personal confidential data from providers without the consent of the data subject could carry other risks. These are based around the potential for patients to lose trust in the confidential nature of the health service. The risk of such a loss of confidence has two facets: first, patients might not receive optimal healthcare if they withhold information from the clinicians that are treating them; and secondly, that this loss of trust could degrade the quality of data for used for commissioning and related purposes, such as risk stratification.

4.2.2 Data collected

In order to evaluate the potential impact on their privacy, patients need to understand what data are to be extracted. Whilst all health data are classified as sensitive personal data under the Data Protection Act, a list of particularly sensitive items will continue to be excluded from extracts. The data extracted is in the format of a series of codes. Free text (i.e., words, sentences, and paragraphs) will not form part of the risk stratification data set to be used.

4.2,3 Benefits from use of data for Risk Stratification

A summary of the benefits, impact and controls of collecting PCD for Risk Stratification is shown in Figure 2.

Figure 2 Collecting PCD for risk stratification - summary of benefits, impact and controls

Reasons for processing and benefits	Impact on privacy	Controls
• The health care activity data collected are fundamental to commissioners understanding 'population health', (e.g. groups of population in need of different types of services, especially specialist or integrated services). • Primary care clinicians are able to contact patients regarded as high risk to take preventative action if the patient wishes. • The management of the	• Some people may feel a loss of individual autonomy (no patient consent) • Some patients not be aware of or understand their choices. • Some patients may be unaware of the use of their data for this purpose.	• Statutory basis for data collection required or permitted by law[7] • Identifiable data must be necessary to satisfy the purpose • Awareness raising activities will help patients understand how their data are used for risk

[7] The Health and Social Care Act provides powers for the Health and Social Care Information Centre to require organisations to submit data to it when data collection has been mandated by NHS England or Secretary of State, and in some circumstances, where requested by other bodies.

health and social care organisations are aided by being able to target resources and service to best effect.		stratification

4.3 Disclosure from General Practice

Under the Data Protection Act, the doctors in a General Practice are the Data Controllers of the data gathered or generated during the course of the provision of services to patients. The General Practice therefore controls the use and processing of the data relating to their patients within the requirements of the Data Protection Act and the common law of confidence and NHS Information Governance policies.

General Practices can use third party data processors to process data under their instruction, (e.g. to provide risk profiles and scores for patients), and can provide de-identified data alongside data from other practices to be used for population based risk stratification. The Section 251 approval for Risk Stratification includes the disclosure of data from GP systems to data processors working under the instruction of GPs as data controllers.

4.4 The processing of the personal confidential data by the HSCIC

Under the Health and Social Care Act 2012, the HSCIC is established as a 'safe haven' with powers to collect and analyse confidential (i.e. identifiable) information about patients. The HSCIC will process the personal confidential data, (e.g. bringing together data from different data sets for secondary care activity submitted to the HSCIC by secondary care providers), to form part of data required for the risk stratification process.

At present, it is not possible for patients to prevent flows of confidential data from other care settings into the HSCIC, for example from hospitals. For this reason, NHS England has ensured that patients can also object (via registering their objection with their GP) to the disclosure of confidential data from the HSCIC, as indicated above in Section 1.13.1.

It is necessary for the HSCIC to receive identifiers so that it can assess data quality and process and link data to form the output datasets to be used in the risk stratification process.

The HSCIC's PIA[8] details the risks and responsibilities it has to protect the confidentiality of all the data it holds, including a large number of datasets containing

[8] http://www.hscic.gov.uk/media/12931/Privacy-Impact-Assessment/pdf/privacy_impact_assessment_2013.pdf

identifiable data. The latter means that the HSCIC are experienced in managing the security and confidentiality of the relevant identifiable data and this is reflected in their PIA.

4.5 The onward disclosure of data from the HSCIC

The relevant HSCIC Data Services for Commissioning Regional Office will

- Disclose the relevant datasets (under the remit of the section 251 application) to relevant authorised Health Service body's controlled environment or accredited third party processor;
- Disclose a weakly pseudonymised data set (with a single identifier) to relevant Stage 1 Accredited ASH or accredited third party processor.

These data are released to the environments with strictly controlled access to PCD or weakly pseudonymised data sets. (ASH).

Figure 3 Safeguards on disclosure of data by HSCIC

The following robust safeguards will be in place in relation to disclosure of data by the HSCIC:

- purpose limitation, (i.e. the data can only be used by the recipient for an agreed purpose or set of purposes);
- training of recipients' staff with access to data, especially on security and data minimisation principles;
- controls over the ability to bring other data into the environment, allowing the risk of re-identification by linkage or association to be managed;
- limitation of the use of the data to a particular project or projects;
- restriction on the disclosure of the data;
- prohibition on any attempt at re-identification and measures for the destruction of any accidentally re-identified personal data;
- arrangements for technical and organisational security, e.g., staff confidentiality terms and conditions of service;
- encryption and key management to restrict access to data;
- limiting the copying of, or the number of copies of the data;
- arrangements for the destruction of the data on completion of the project; and
- penalties, such as contractual ones that can be imposed on the recipients if they breach the conditions placed on them.

Whilst there is privacy risk that the analysts granted access to these pseudonymised flow could potentially re-identify patients maliciously by combining the pseudonymised data with other available datasets (a technique known as a jigsaw attack) such an attack would be illegal and would be subject to sanction by the ICO and enforcement action by the Department of Health under the Section 251 regulations.

As stated in section 1.15, there is not a straightforward process for patients to prevent data flows from other care settings, e.g. hospitals, to the HSCIC. However patients can register their objection with their GP practice to prevent personal confidential data derived from any healthcare setting leaving the HSCIC unless there is an overriding public interest such as a civil emergency.

Figure 4 HSCIC Processing of PCD for subsequent use in Risk Stratification - summary of benefits, impact and controls

Reason for processing and benefits	Impact on privacy	Controls
• Statutory basis for collection and analysis. • Processing within HSCIC, the NHS's major Save Haven • Accuracy has to be checked before data are de-identified (it is not possible afterwards)	• In some cases, a small residual risk that identifiable data could be revealed • Risks of jigsaw attacks increase as more effectively anonymised data are made available, to more organisations. • Data collection, storage and processing creates risk of confidential information being accessed without the knowledge or consent of patient • Risks in terms of changes to scope (e.g. to dataset or use) without patients being aware.	***Potentially identifiable data:*** • Robust information governance controls will be applied as detailed in **Error! Reference source not found.**. ***Personal confidential data:*** • Patients can object to their personal confidential data leaving the HSCIC. • Identifiable data stored only where necessary and destroyed or aggregated, anonymised or pseudonymised as soon as possible. • Patient identifiers are held separately from clinical data within the HSCIC. • De-identifying data reduces or eliminates the risk of a person's identity being revealed and thus helps protect privacy. • Contractual provisions and oversight to ensure data are only used for RS purposes or for other legitimate purposes as agreed under contract.

4.6 Processing of data for Risk Stratification

Processing of data for risk stratification takes place under the constraints set in place by the approval of the Section 251 by the Secretary of State. This means that processing can only be undertaken by accredited organisations, either already under contract to the NHS with a proven track record on managing data for risk stratification or by Commissioning Support Units, effectively part of NHS England, that have achieved (Stage 1) ASH status.

The data for risk stratification and the related processing are held independently (or through virtual separation) to prevent use with other data.

As indicated in Figure 1, the first step in processing is to pseudonymise the received data. Thus the processing for risk stratification is undertaken with a weakly pseudonymised data set (with only one identifier) or with a fully pseudonymised data set. Using either of these data sets would ensure that the identity of the individual is unknown, as the means are not available to re-identify from the full or weak pseudonymisation.

Initially, more data than the minimum amount of information necessary may be provided by general practice and from the HSCIC. This is because the extraction facilities from general practice systems may be generic and produce a complete record of coded data, but only relevant data items will be processed for risk stratification and the remainder removed and destroyed.

As indicated earlier, see section 1.8, risk stratification involves the linking of data from general practice with data from secondary care activity. This should take place through linking the pseudonymised data, so that identifiable linked data are not visible during any processing steps.

Data reflecting episodes of care by general practice and by secondary care organisations for a period, say three years, for all relevant individuals within an area, (e.g., CCG), are linked for each individual. Statistical analyses are then applied to the linked records through computational algorithms on factors, such as age, sex, diagnosis, treatment code or length of stay in hospital. These algorithms may be published or may be proprietary and maybe locally modified to reflect specific local interests. These algorithms produce a score for each individual patient in that area.

Staff in the relevant accredited organisations will have contracts restricting their access to identifiable data to their roles, (e.g. database administrator resolving processing problems for the initial pseudonymisation).

The further actions include:

- Fair Processing Notices by commissioners and General Practice
- Ensuring that the right to object arising both from Section 10 of the Data Protection Act and as set out in the NHS constitution and by policy are considered and captured within systems
- Detail how objections expressed at one point (for example, with the GP or with the HSCIC) will be taken account throughout the system.

4.6.1 Risk Stratification for commissioning

The output sought by commissioners is along the lines of risk stratifying the whole population in order to identify those most at risk of a hospital admission in next 12 months and planning for services to support their care. This can be achieved through categorising the scored records generated as described in section 1.17 above in different ways, such as by area, condition (e.g. for diabetes) or by age group. The relative volume of records or relative scores enable a picture of the health and needs of the population to be developed, in effect a form of health needs assessment – a key requirement for commissioners to meet their statutory obligations. In turn this can enable priorities to be determined in the use of resources and planning services. The output required by commissioners is therefore in the form of aggregate results or possibly in some instances at individual patient level, but without the need for any patient to be identified. The latter use is for local analysis and to cover the range of potential questions and issues that Commissioners may need to consider.

4.6.2 Risk stratification for case finding

The purpose of risk stratification for clinicians in general practice is to be able to be aware of only those patients who are likely to need hospital or other healthcare services in order that the patients can be approached about a suitable intervention and their consent and participation sought in follow up to provide the intervention.

To meet this purpose, it is necessary to identify relevant patients to approach them. Authorised clinicians, often community matrons, are usually responsible for handling the sensitive information about an identified individual and making appropriate approaches to them. To do this, a different output is required from that mentioned in section 1.17.1, namely access to records and scores for identified individuals. Access to such data are controlled through role based access controls (RBAC) to relevant files/portal from the risk stratification system. The provision of identifiable data should be carried out through linking of relevant pseudonymised records with patient identifiers are held separately from clinical data in production of the report for the authorised end user.

4.7 Conclusion of privacy issues as a consequence of risk stratification

The main tension identified within this privacy assessment is the balance between the

- benefits of
  - A) using linked de-identified clinical data from health services to improve health needs assessment, service planning quality, in order to improve commissioning of services with a focus on need, outcomes and patient experience;
  - B) using linked clinical data from health services to provide a relative high scoring risk assessment, which is identifiable only by authorised clinicians for only those individual patients who need to be approached in order to provide relevant services to provide potentially improved outcomes, quality of life and patient experience

- the risks to patient privacy from the collection, linkage, analysis, storage and disclosure of the data
  - either in de-identified form to a restricted audience of commissioning staff
  - or in identified form to a restricted set of authorised individuals involved in a direct care relationship to the patient.

A summary of benefits and privacy issues is shown in Figure 5.

Figure 5 Processing of PCD in Risk Stratification - summary of benefits, impact and controls

Reason for processing and benefits	Impact on privacy	Controls
• To link data to develop longitudinal records for analysis • To provide risk assessment scoring on individual patients in de-identified form • To develop population risk profiles from aggregation of individual scores and analysis into subgroups, such as areas, conditions, age groups • To provide risk assessment scoring on individual patients in identifiable form • To develop cohort lists for general practice of patients at high risk of need of care services, so that services can be appropriately offered	• In some cases, a small residual risk that identifiable data could be revealed • Risks of jigsaw attacks increase as more effectively anonymised data are made available, to more organisations. • Data collection, storage and processing creates risk of confidential information being accessed without the knowledge or consent of patient • Risks in terms of changes to scope (e.g. to dataset or uses) without patients being aware.	***Potentially identifiable data:*** • Robust information governance controls will be applied as detailed in **Error! Reference source not found.**. ***Personal confidential data:*** • Patients can object to their personal confidential data leaving their practice or the HSCIC – public awareness supported by privacy notices (see Section **Error! Reference source not found.**). • Pseudonymisation on landing separates patient identifiers from clinical data. • De-identifying data and de-identified data for commissioners reduces or eliminates the risk of a person's identity being revealed and thus helps protect privacy • Patient identifiers are only available to authorised clinical users within general practice. • Contractual provisions and oversight to ensure data are only used for RS purposes or for other legitimate purposes as agreed under contract.

A key component of any assessment is the degree to which these risks are mitigated by the controls and security that will be applied.

Data will be transferred in encrypted identifiable form, but will be pseudonymised on landing, processed in pseudonymised form, output in aggregate or pseudonymised form for commissioning purposes and only be revealed in identifiable form to authorised users to minimise the risks to an individual that their privacy will be breached.

The potential risks to privacy from risk stratification processing are:

- Loss of individual autonomy from use of patient identifiable data without consent
- Risk of confidential information being accessed and viewed without knowledge or consent of patients
- Linking and de-identification processes may not be reliable enough to achieve total anonymisation of data
- Risk of data being accessed illegally and then sold or otherwise misused by commercial organisations, criminals or others
- Risk of data being accessed legally and then the data being misused.
- The actual mitigating controls required under Section 251 of the Health and Social Care Act of 2006 and its supporting regulations, and NHS best practice in relation to third party processors will use to safeguard these risks are summarised below.

The risk stratification process will require accredited processors to:

- Obtain and process only the minimum necessary patient identifiable data from other organisations
- Store and process data in their accredited safe haven or equivalent
- Keep to the absolute minimum the number of staff able to access and view patient identifiable data, and wherever practicable assign staff rights of access to either patient identifiers or clinical data but not both
- Destroy data held in identifiable form as soon as they are no longer required, or in accordance with the relevant NHS retention policy
- Disclose only anonymised or aggregate data, unless there is a legal basis for the disclosure of confidential data to authorised users
- When disclosing anonymised data, restrict the data disclosed according to the context in which the data will be used.
- Monitor who accesses patient identifiable data.

5. Alternatives to identifiable data

Identifiable data are always likely to be needed for risk stratification purposes, unless a NHS corporate wide pseudonymisation facility is introduced. The need for identifiable data is two fold – one to enable data from disparate primary and secondary care sources to be linked prior to risk scoring, and secondly the need to output identifiable data to authorised clinicians for contacting high-risk patients that are to have their needs assessed and additional care offered.

As indicated in section 1.6, it is expected that the gaining of identifiable data and its processing will be undertaken under the auspices of the HSCIC and its legal framework unless this proves not feasible.

A relevant privacy impact assessment will need carrying out at that time, but it is expected that the existing HSCIC PIA will cover the risk stratification.

6. What will we do to protect privacy?

6.1 Overview

The HSCIC and their DSCROs undertake part of this process. Their core purpose within legislation is to process patient records safely and securely. As stated in the HSCIC PIA "The HSCIC has been processing patient records safely and securely since its inception. It has introduced strong security controls, published and implemented security policies and published information about its processing as required for compliance with the Department of Health's Information Governance Framework. … The HSCIC takes its responsibilities as a custodian of patient information extremely seriously".

It is hoped that the managed change programme will enable the HSCIC to undertake all processing where identifiable data are required and that only pseudonymised or weakly pseudonymised data are used within the commissioner. However, this is subject to overall NHS England's 'Defining and delivering the data required for commissioners' work and its consent strategy.

6.2 IG Controls within CSUs and authorised third party data processors

The requirements on data processors for information governance controls will reflect information governance commitments made by the HSCIC. The CSUs and third party data processors will have to achieve standards such as

- satisfactory completion of the NHS Information Governance Toolkit Level 2
- compliance with ISO27001/2 Information Security Standards, which include:
  - Staff training and contracts?
  - Information technology system security and audit trails
  - Robust management arrangements?
  - Full compliance with legislative requirements
  - Provision of the "safe haven" for sensitive information.

6.3 Local risk stratification privacy impact assessments

This PIA for risk stratification is necessarily a high level and generic PIA to support the Section 251 of the Health and Social Care Act 2006 approval. In order that the specific risks that may arise within particular local operation of risk stratification data processing, it will be a compliance requirement on organisations responsible for such processing to undertake a similar PIA risk assessment.

7. Public acceptability

The HSCIC PIA describes how the government consulted stakeholders to inform the powers in the Health and Social Care Act[9].

Personal confidential data have been used for purposes beyond direct care for many years such as for healthcare planning and for research. Risk stratification is another example of where benefits can accrue from the use of patient data for indirect care purposes, both for the development of healthcare services and for the benefit of individual patients at risk.

It is important that patients are clear about what information is being shared, how it is being shared and why so that they can understand the risks and benefits to them and to the wider population. Some patients may have particular concerns and therefore NHS England has made it simple for patients to object.

7.1 Independent scrutiny

Scrutiny to the use of PCD for the purpose of risk stratification was provided by the Confidentiality Advisory Group (CAG), part of the HRA, as part of the application for section 251 support setting aside of the Common Law of Confidence.

The ICO provides expert advice to the CAG on compliance with the Data Protection Act. Under section 251 whatever is done in the setting aside of the Common Law of Confidence for obtaining consent to process data, the standards and controls associated with the Common Law of Confidence still apply as does the need to be compliant with the Data Protection Act.

7.2 Patient Information Materials

Patient information materials will be produced at national and local levels. The requirements for such materials, topics to be covered etc., will be outlined in the forthcoming Fair Processing Strategy by NHS England.

8. Conclusions

Any processing or storage of identifiable patient data introduces potential risks of data misuse and breaches of privacy. Although they can never be eliminated, such potential risks are significantly mitigated by the robust information governance controls as set out in sections 1.11 and 1.20 which are all designed to safeguard patients' privacy.

However, the processing of a person's information without their permission is a loss of autonomy for that individual. For this reason, in addition to the extensive safeguards for the data, NHS England is supporting data controllers to raise awareness among patients and making it simple for patients to object to the disclosure of personal confidential data.

[9] HSCIC PIA section 2.3 (http://www.hscic.gov.uk/media/12931/Privacy-Impact-Assessment/pdf/privacy_impact_assessment_2013.pdf

In summary, people who conclude that the net impact of risk stratification on privacy will be positive are very likely to be supportive of the CCGs use of risk stratification. Even people who feel the impact will be detrimental to privacy may recognise that the potential benefits of risk stratification using data from patient records are great, and may therefore feel they are justified ethically on that basis. However, some people may believe that any use of patient identifiable data without explicit patient consent is unacceptable. These people are unlikely to be supportive of risk stratification whatever its potential benefits and may object to the use of personal confidential data for wider healthcare purposes.

9. Abbreviations used in this paper

ASH	Accredited Safe Haven
CAG	Confidentiality Advisory Group
CCG	Clinical Commissioning Group
CSU	Commissioning Support Unit
DPA	Data Protection Act
DSCRO	Data Services for Commissioners Regional Offices
HRA	Health Research Authority
HSCA	Health and Social Care Act 2012
HSCIC	Health & Social Care Information Centre
ICO	Information Commissioner's Office
ICT	Information and Communications Technology
PCD	personal confidential data
PIA	Privacy Impact Assessment
RBAC	Role Based Access Control
RS	Risk Stratification