NHS

National Programme for
Information Technology

Restricted - Policy						
Programme		DOCUMENT NUMBER				
Sub-Prog/Project	SUS	National Prog	Org	Prog /Proj	Doc Type	Seq
Prog. Director	J. Thorp	NPFIT	NCR	DES		
Sub Prog/Proj Mgr	L Franklin					
Author	JF & WG	Version No		0.8		
Version Date	8/12/05	Status				

This document is provided for information, although comments are very welcome, particularly in relation to Shared Services.

We will be separately progressing with CRDB as well as presenting to SCAG and contacting PIAG.

Jeremy Thorp

19 January 2006

A Framework for SUS Information Governance

Secondary Uses Service

16 January 2006

Restricted

NHS

National Programme for Information Technology

Restricted - Policy						
Programme		DOCUMENT NUMBER				
Sub-Prog/Project	SUS	National Prog	Org	Prog /Proj	Doc Type	Seq
Prog. Director	J. Thorp	NPFIT	NCR	DES		
Sub Prog/Proj Mgr	L Franklin					
Author	JF & WG	Version No		0.8		
Version Date	8/12/05	Status				

Document control sheet

Client	Secondary Uses Service
Document Title	A Framework for SUS Information Governance
Version	0.91
Status	
Reference	
Author	Jonathan Fistein & Wally Gowing
Date	16 January 2006

Document history			
Version	Date	Author	Comments
0.91	16 January 2006	Jonathan Fistein & Wally Gowing	

Approvals

This document requires the following approvals.

Name	Signature	Title	Date of Issue	Version
J Thorp				

Document Location

This document is only valid on the day it was printed. Please contact the Document Controller for location details or printing problems.

This is ~~a controlled~~ an uncontrolled document.

On receipt of a new version, please destroy all previous versions (unless a specified earlier version is in use throughout the project).

Contents

1 Executive Summary and Key Recommendations

1.1 Introduction

1.1.1 This document argues that there must be an effective Information Governance framework for SUS that balances the need to protect patient confidentiality against the need to share patient data to support legitimate healthcare business processes.

1.1.2 As SUS develops it will give its users sophisticated tools that will increasingly support their business processes while greatly reducing their need to access identifiable data (as highlighted in the Pseudonymisation Impact Assessment Study (PIAS)[1] and shown schematically in Figure ES1).

1.1.3 In the short term, the use of clear data will be moderately prevalent to support legitimate NHS processes, and it is particularly important to develop robust and defensible information governance procedures to ensure there is no inappropriate access to identifiable data. Additionally there is an urgent need for procedures to support shared service activities within SUS, as these are an essential component of NWCS decommissioning. Therefore this document proposes a number of interim information governance measures for shared services and for short-term access to clear data that are informed by and aligned with the principles behind the guidance issued by the Care Record Development Board (CRDB), Hospital Episode Statistics (HES), Office of National Statistics (ONS), Connecting for Health (CfH) Digital Policy Team and others.

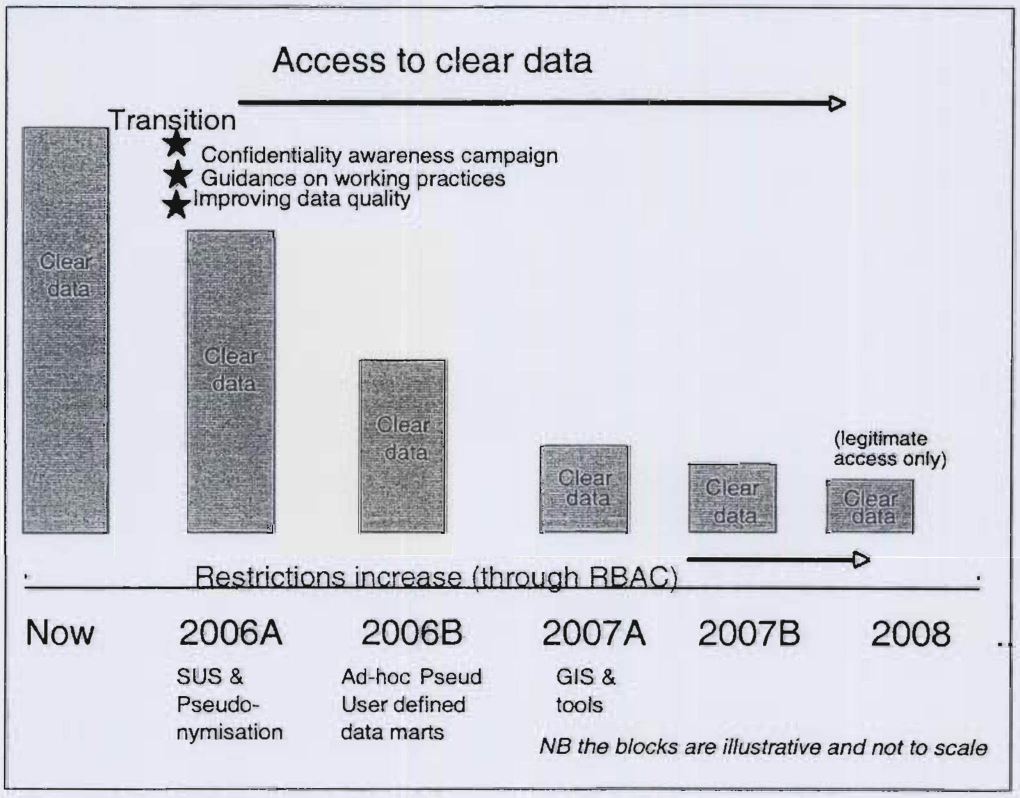
1.1.4 Also described are frameworks to support the transitional period before the full introduction of pseudonymisation (i.e. before clear data flows are replaced by pseudonymised data flows) and a SUS Protocol to allow users access to the data they require for their legitimate business processes without any disproportionate technical or procedural overheads.

1.1.5 It is recognised that SUS Information Governance rules must sit within the wider organisational context of CfH and the NHS. There are several overlapping technical domains that are expected to provide secure environments to help ensure patient confidentiality, including Spine initiatives such as Spine Directory Services (SDS) and Role Based Access Control (RBAC). Although these all are constrained by the broad policies contained within the Care Record Guarantee and the Confidentiality Code of Practice they are each developing specific rules for their own information governance which can directly or indirectly affect SUS. It is recommended that SUS liaises closely with these other initiatives to ensure that the overall system enables SUS users to perform their legitimate business functions in a way that is compatible with good information governance practice.

[1] Secondary Uses Service, Pseudonymisation Impact Assessment Study 17/09/2005 (Paper to the SUS Project Board)

1.1.6 Finally, it is argued that an appropriate independent body is formed that can assume responsibility for Information Governance for SUS. This should have the power to both set the overall policy direction for SUS and to take decisions about whether individual requests for access to data are appropriate.

Figure ES1-The decreasing use of clear data in SUS



1.2 Recommendations

1.2.1 In the short term, it is vital that NHS organisations are granted the access they require to support their essential business processes. During this period, it is likely that a number of information governance issues will arise and a means of handling the escalation of these is required.

1.2.2 *Recommendation 1* - It is recommended that consideration be given to how such issues are handled and, as discussed in Section 4, to the appropriate mechanisms to manage Information Governance within SUS.

1.2.3 For the immediate future, it is recommended that the following are discussed by the SUS Board and then put forward as proposals to the CRDB:

- *Recommendation 2* - The Transitional arrangements as described in Section 5. These include:
  - handling clear data extracts
  - distinguishing users' roles.
- *Recommendation 3* - The SUS Protocol approach to the management and use of data extracted from SUS and described in Section 5.3
  - the aims
  - an initial outline Protocol.
- *Recommendation 4* - as described in Section 6, the approach to enable shared services to access SUS whereby:
  - A clear legal framework for shared services is established
  - Organisational relationships in shared services are elucidated
  - "Lead" organisations in shared services register the shared service with NACS
  - The technical infrastructure is developed to enable shared services, starting with the proposed short-term fix using the CRISP component of the Address Grid
  - A work plan is established with the National Application Service Provider (NASP) to provide shared services functionality to support NWCS Decommissioning against the timetable in Section 6.6.2 in order to be operational before June 2006.
  - The NASP is requested to produce proposals for more strategic solutions to supporting Shared Services as set out in Section 6.4.3

1.2.4 *Recommendation 5* – As set out in Section 4, it is recommended that, working within the overall CRDB and CfH information governance guidelines, an Information Governance Body for SUS is established to:

- Direct SUS Information Governance Policy
- Establish clear procedures for enacting the Information Governance (IG) Policies
- Liaise with CfH Digital Policy Team, NHS and other Information Governance bodies
- Establish appropriate frameworks for non-NHS access to SUS data

1.2.5 *Recommendation 6* - In the medium to long term, it is recommended that this body takes an approach as described in Section 2 that:

- Recognises that there must be an appropriate balance between the competing public goods of protecting confidentiality and using patient data to improve healthcare.

- Stresses the different implications for information governance of sharing data within the controlled, secure SUS environment versus releasing information from that environment.

1.2.6 *Recommendation 7* - As described in Section 3, it is recommended that further work be commissioned to meet unresolved technical and policy challenges associated with SUS surrounding:

- Small number handling
- The application and governance of RBAC
- The formulation of rules for the provision of clear, pseudonymised and aggregate data for different purposes
- Stop-notes and patient consent
- The use of aggregate data and derived fields to protect confidentiality.

2 An Approach to Information Governance for SUS

2.1 Introduction

2.1.1 SUS will contain identifiable patient data and it is essential that this is held within a secure environment with appropriate access rules, in line with best practice for information governance. In the medium to long term these rules should be formulated and enforced by an appropriate (independent) body as described in Section 4. However, there are several practical issues facing SUS for which polices and rules for Information Governance are required in the immediate future. These issues include arrangements for shared information services and the transitional arrangements before clear data flows are replaced by pseudonymised data flows.

2.1.2 It is recommended that an interim, pragmatic approach to Information Governance in SUS is taken to enable the data flows that are required for the business continuity of SUS users in the short-term during NWCS decommissioning. This is informed by:

- Current DH policies and guidance, e.g. Care Record Guarantee, *Confidentiality*[2]
- Current legislation, e.g. Data Protection Act (DPA)
- The current technical infrastructure within CfH e.g. RBAC
- Organisational constraints of the NHS
- The needs of users, who are using data to support essential business processes

2.1.3 It is recommended that a balanced approach is taken to ensure that "reasonable and proportionate" efforts are used to protect patient confidentiality while still allowing access to data to support legitimate healthcare business processes. This approach could form the framework for the longer-term Information Governance policies and procedures for SUS.

A multi-layered approach

2.1.4 SUS will include a range of technical solutions that will protect patient confidentiality by controlling access to sensitive data. However it must be recognised that technical solutions can only be the starting point for a robust system of information governance. Any technical solution is embedded within its organisational context, and it is therefore necessary to have a "chained" or multilayered set of policies and procedures. For SUS these layers are:

- **SUS itself**, providing technical solutions that will provide a secure environment to protect patient confidentiality (e.g. separation of data into marts, differential access to data based on user role, pseudonymisation, derived data, small number handling, linkage engines, etc.). These should be constrained by appropriate information governance rules.

[2] NHS Code of Practice for Confidentiality

- **CfH**, which provides infrastructure (such as secure physical environments, NHS National Network (N3), Spine), enabling technologies (such as RBAC, SDS) and policy guidance (from the CfH Digital Policy Team)
- **The wider NHS**, which constrains its employees with policies and rules relating to Information Governance, and which should adopt working practices that are consistent with the principles of good information governance.

2.1.5 Technical solutions (together with policies for their use and maintenance) therefore are a necessary part of the wider system of information governance but they are not in themselves sufficient to guarantee confidentiality, as human factors must also be taken into account. Ultimately some NHS staff will legitimately require access to identifiable data and there must be some level of trust in these individuals to use this data responsibly and appropriately.

2.1.6 In this context it is important to recognise the distinction between:

- **Sharing information within a trusted environment, for example within the NHS or between the NHS and other trusted third party organisations or individuals.** Trusted environments are characterised by their level of control i.e. there are technical and organisational constraints on the flow and uses of data and on individuals. These controls may be technical (e.g. having separate data marts that do not allow the mixing of identifiable and anonymised data) or may be contractual (e.g. imposing conditions on data usage on NHS employees; having contractual arrangements with trusted third parties that make their obligations and responsibilities clear); and
- **Publishing information to an insecure environment.** In this case information is released from within the trusted environment and there is subsequently little or no control over the usage of that data, for example when aggregate data is released into the public domain.

2.1.7 This distinction can be mapped on the "purpose of use" approach adopted in *Confidentiality*; however the approach taken in *Confidentiality* largely examines requests for access to data on an individual case-by-case approach. It is proposed that the possibility of expanding this to include policies and procedures terms of "classes of use" should be explored, so that guidance about how grouped data in SUS should be accessed appropriately.

2.2 Balancing Risks

Anonymisation vs. Utility

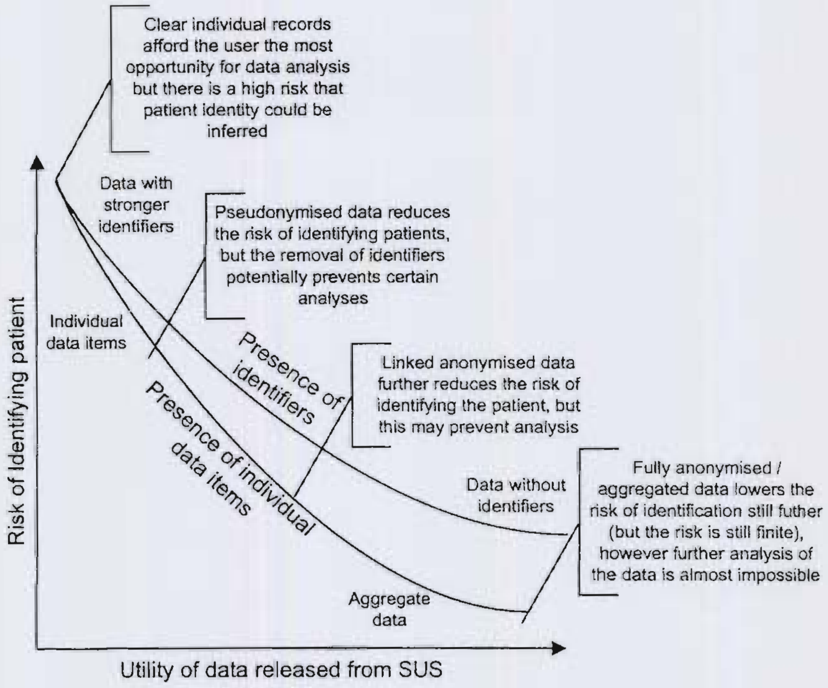
2.2.1 SUS will use a number of technologies to reduce the ability to identify patient data before it is used. It should be recognised that these have important implications for the utility of the information. As a general rule, the more a data set is de-identified and / or aggregated, the less is the utility of the information (see Figure 1). This is illustrated by examples at either extreme: data sets that contain complete patient data offer the best opportunities for analysis, de-duplication and linkage, whereas in contrast it would be impossible to cross-link two aggregated, anonymised data sets. It is the overall vision of SUS to reduce end-users' requirements for clear data by providing tools that will enable them to perform analyses within the SUS (for example using pseudonymised or aggregate data).

2.2.2 Therefore an appropriate balance must be struck between:

- Providing identifiable data, with the risk of breaching confidentiality; versus
- Providing some form of anonymised data, with the risk that meaningful analysis of this data is difficult or impossible

2.2.3 Policies are required that specify which kinds of data are appropriate for different purposes. General rules should be formulated to encompass the majority of data uses; however procedures will also be required to deal with any exceptional requests for data.

Figure 1 - Schematic diagram showing the relationship between Ability to identify Data and perceived Utility


Clear individual records afford the user the most opportunity for data analysis but there is a high risk that patient identity could be inferred
Data with stronger identifiers
Pseudonymised data reduces the risk of identifying patients, but the removal of identifiers potentially prevents certain analyses
Individual data items
Presence of identifiers
Linked anonymised data further reduces the risk of identifying the patient, but this may prevent analysis
Presence of individual data items
Data without identifiers
Fully anonymised / aggregated data lowers the risk of identification still futher (but the risk is still finite), however further analysis of the data is almost impossible
Aggregate data
Risk of Identifying patient
Utility of data released from SUS

Number of users able to access data vs. risk of identifying patients from the data

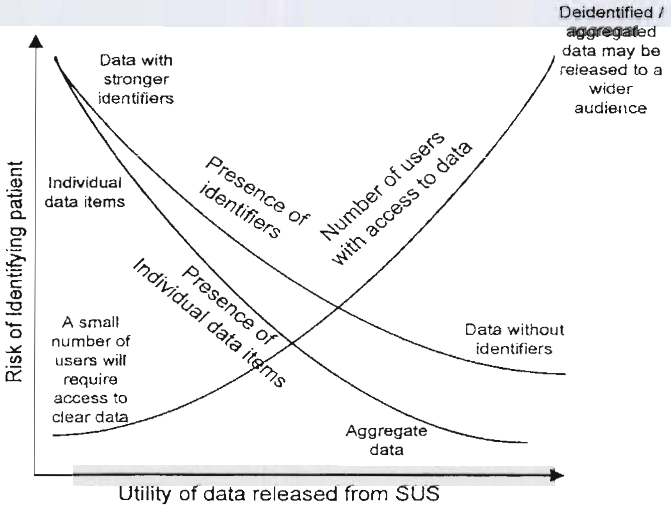
2.2.4 As discussed above, some users *will* require access to identifiable data in order to perform their contracted and statutory duties. As outlined in the Pseudonymisation Impact Assessment Study, both the number of individuals requiring such access and the amount of clear data required will diminish with successive releases of SUS as more internal functionality is developed. However, a small number of users will require access to clear data in the long term. Further study into these is required to determine the circumstances in which clear data access is permitted and the legal basis for this.

2.2.5 The risk to compromising the confidentiality of patient data is a function of:

- The sensitivity of the data
- The ability to identify the subject of the data
- The number and trustworthiness of those accessing the data

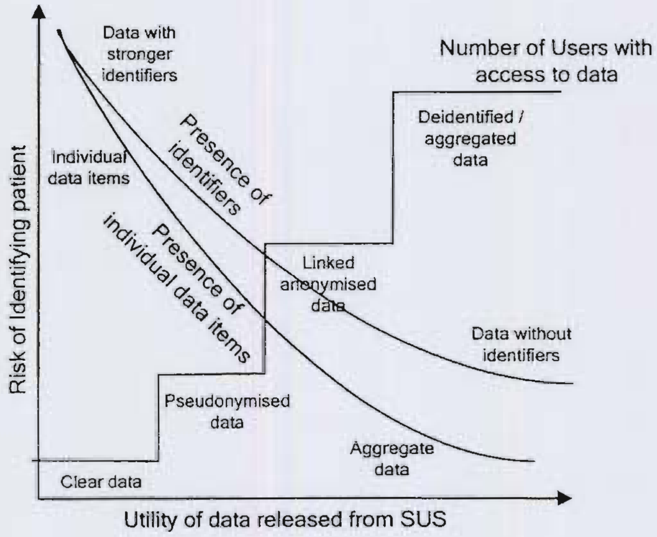
2.2.6 The access control mechanisms within SUS will ensure that only appropriate users can access identifiable data directly. For the majority of users and users aggregate and / or pseudonymised data will be provided that will greatly reduce the risk of identifying individuals. This is illustrated in Figure 2.

Figure 2 - Schematic diagram showing numbers of people able to access data on SUS


Deidentified / aggregated data may be released to a wider audience
Data with stronger identifiers
Individual data items
Presence of identifiers
Number of users with access to data
Presence of Individual data items
A small number of users will require access to clear data
Data without identifiers
Aggregate data
Risk of Identifying patient
Utility of data released from SUS

2.2.7 In the figures above, the de-identification of data is represented as a continuum. It may be more practical, however, to separate this into categories as shown in Figure 3 (only four broad categories are illustrated for clarity). Policies and procedures for allowing access to data in each of these categories will be required. These should take into account the broader principles of Information Governance, including the purpose of use, the sensitivity of the data, the size of the audience, etc.

Figure 3 - Schematic diagram illustrating a categorised model for access to data on SUS


Data with stronger identifiers
Number of Users with access to data
Deidentified / aggregated data
Individual data items
Presence of identifiers
Presence of individual data items
Linked anonymised data
Data without identifiers
Pseudonymised data
Aggregate data
Clear data
Risk of Identifying patient
Utility of data released from SUS

Risk of identifying patients from data vs. the effort it takes to de-identify the data

2.2.8 Techniques to aggregate and de-identify data have costs, both in terms of the reduction in utility of the data that is ultimately provided as described above, but also on terms of the effort it takes to develop and apply the techniques.

2.2.9 As discussed above, the sensitivity of the data must be taken into account when considering the appropriate amount of effort that should be expended to prevent the potential identification of patients. Extra effort may be required to ensure patient identifiers can not be inferred when dealing with data in particularly sensitive clinical domains. For example, currently IVF records are anonymised on output from trust's systems and currently go through NWCS in anonymised form. There are similar sensitivities surrounding communicable disease and terminations data.

2.2.10 Several bodies have looked at these in detail and it is recommended that further work is commissioned to devise a detailed policy for SUS that is in line with current best practice.

3 Components

3.1 Technical controls within SUS

3.1.1 As discussed above there are a number of technical components within the SUS architecture that can support sound Information Governance. These, with their associated governance procedures, can help to ensure that the confidentiality of records is protected appropriately. The mechanisms for these should be developed and then publicised in order to reassure patients and third parties who submit data to SUS (such as the ONS) that this data is secure.

The use of RBAC

3.1.2 Access to data within SUS will be controlled by means of RBAC. In the RBAC system, users are presented with a set of User Role Profiles (URPs) each of which represents a particular purpose of use for which a user may require access to SUS data. This is represented in the URP as the Organisation on behalf of which the user is acting and the activities (business functions) that the user is allowed to perform within that role.

3.1.3 Users will be required to select one URP at SUS logon. The SUS application will then present functionality and data appropriate to only that URP. This will be particularly important during the interim (2006-A and B) period before all of the pseudonymisation requirements have been implemented as users may potentially have access to both clear and pseudonymised data about the same patients from within different roles. It is essential to keep reports based on these separate to prevent inappropriate identification of patients.

3.1.4 In order for this to be possible, the following rules will be applied:

- Application functionality within SUS will not allow reports to contain both clear and pseudonymised data about the same patients
- Business functions will not allow access to different applications functions that could potentially allow users to cross-reference pseudonymised and clear data about the same patient.
- URPs will not be allowed to contain combinations of Business Functions that could potentially allow users access to both pseudonymised and clear data about the same patient.

3.1.5 It is anticipated that in the long term, most users will only have one URP that allows them access to SUS, as most of their tasks will be supported by internal SUS application functionality and pseudonymised data. However, some users will require access to both clear and pseudonymised data as they carry out different roles at different times. This will necessarily require the users to have multiple URPs to allow differential access to the SUS data. In such cases it will be technically impossible to prevent users from downloading clear and pseudonymised data about the same patient in different user sessions. The impact any potential abuses of this can be minimised by either not allowing users to download (clear) data for offline analysis or by insisting that users store clear and pseudonymised data separately. However both of these approaches may have implications for essential business processes that will need further investigation.

Pseudonymisation

3.1.6 A set of rules needs to be developed to control the application of pseudonymisation to data held on SUS. These should make it clear when it is appropriate to provide data either in the clear or having undergone pseudonymisation. These rules will necessarily have to be detailed, going down to the field level, and will have to take into account potentially numerous purposes of use for data items.

3.1.7 In parallel with these, there must be rules for the generation, maintenance and control of the pseudonymisation keys used for generating the pseudonyms.

Small number handling

3.1.8 Data sets may enable the identification of individuals by:

- Including identifiers
- Implying the identity of individuals, for example by returning results that contain small numbers of patients in restricted topics or geographic areas

3.1.9 It is important to recognise that removing identifiers alone is not enough to necessarily prevent the identification of patients, but other techniques will be required to protect confidentiality. Rules will be required in order to reduce the possibility of the identification of patients from de-identified results that contain small numbers contained within small geographic areas or within sensitive clinical areas. SUS should develop a strategy to reduce the risk of identification of patients from "de-identified" data where appropriate. The development of this strategy should be informed by the previous work that has been performed by other centres including the ONS, HES, UK Association of Cancer Registries (UKACR), etc. It should be recognised that small number handling is complex as a variety of factors need to be taken into account.

Aggregation of data and presentation of derived fields

3.1.10 The needs of a large number of data users can be met by providing aggregate data or by providing derived information from the primary data e.g. providing ages rather dates of birth. It is envisaged that if these derived and aggregate data are provided from within SUS, there will be a large reduction in the number of requests for clear data from users. Guidelines are required to specify which derivations should be provided for different purposes. It is proposed that these are informed by existing guidelines from HES and others.

Data quality considerations

3.1.11 Currently many users request clear data as they feel they have to see the primary source data in order to solve data quality issues. These include correcting missing data fields, removing duplicates, and standardising data formats. SUS will correct the majority of these errors on import. It is proposed that an external facing "data quality mark" is created to further reassure users that they can rely on the (non-identifiable) data within SUS.

3.1.12 As part of the data quality initiative, redundant data should be removed. One example is the need to remove the use of Very General Purpose (VGP) fields as these could contain patient identifiable data and could compromise patient confidentiality. VGP fields were allowed in the original activity datasets to enable information to be exchanged between providers and commissioners on a local and informal basis. There are now other means of providing such information and VGP fields are therefore redundant.

"Stop-notes" and patients withdrawing consent for the transmission of data to SUS

3.1.13 According to the principles outlined in *Confidentiality*, patients may have the right to refuse to allow their data to be made available on SUS. This potentially seriously limits the usefulness of the data contained within SUS. For example, certain demographic groups may preferentially opt to withhold their data from SUS, which would introduce a bias into the data and confound later analysis. Such withholding of data may have similar effects to confound analyses of healthcare data on SUS for healthcare management purposes.

3.1.14 A similar situation arises in the case of "stop-notes" i.e. where data is either not imported into SUS or is not presented on export because it relates to a celebrity or other VIP, however the numbers of such VIPs is likely to be relatively small.

3.1.15 Further clarification of the policies relating to patients not consenting to having their data on SUS and analysis of the operational implications are required.

3.2 Connecting for Health Policies

RBAC Governance

3.2.1 RBAC is the "key" that allows users access to SUS data. It is therefore essential that there is adequate control over the allocation of RBAC roles to ensure that keys are only given to appropriate users. This control is outside of SUS, falling to the Registration Authorities (RA) that create the RBAC User Role Profiles. It is currently possible for any Registration Authority to create User Role Profiles that contain any Business Function and any Organisation Code. This potentially represents a serious threat to the confidentiality of the data held in SUS.

3.2.2 Historically, URPs have been created that contain inappropriate organisation codes and/or inappropriate Business Functions.

3.2.3 It is understood that in future releases of RBAC there will be controls over the RAs to:

- Restrict which Business Functions can be allocated at local RA level; this is essential to ensure, for example, that Business Functions that provide access to clear data are not allocated inappropriately.

- Restrict the organisations the RA can put into the URPs it creates to the RA itself or its subsidiary organisations

Additionally there will be an RBAC Governance Board formed outside SUS. One or more members of the SUS Project Board will be invited to report to the RBAC Board.

3.2.4 A strategy is required to deal with the past incorrect allocations of organisations and Business Functions in order to ensure that there is no inappropriate access to data. A review of the current RBAC arrangements has shown that refinements will need to be made to the current SUS Business Functions. This could present a useful opportunity to "retire" some of the Business Functions that have been inappropriately allocated to users, which could be replaced by Business Functions under closer control.

3.2.5 In the short term, it is recommended that guidance be issued to RAs on acceptable use of SUS Business Functions. This usage can be retrospectively monitored via the spine by examining the URPs that have been created. However, it must be recognised that in the absence of formal controls, inappropriate access to data may still be possible.

3.2.6 Going forward, it is recommended that SUS uses its influence on the RBAC Governance Board to ensure that there are effective controls over the Registration Authorities and that there are clear procedures and guidelines for:

- **Defining (and maintaining) the sets of baseline Business Functions associated with roles in the RBAC dictionary,** for example to ensure that Business Functions are not combined that would allow access to the same data in both pseudonymised and clear forms.
- **Defining acceptable uses for Business Functions**, for example defining to which roles Business Functions may be allocated as "extras", and defining the procedures for allocating extra Business Functions to URPs.
- **Defining an appropriate and workable framework within RBAC to support Shared Services.** There are a number of settings within the NHS where one 'lead" organisation may be authorised to access the (clear) data of one or more other organisations. In these cases it is necessary to carefully control and audit access to ensure this cross-organisational access is only provided to appropriate individuals. Examples include joint informatics services, joint commissioning arrangements and joint service providers. The situation is further complicated as a single organisation may "lead" different joint commissioning arrangements, for example in different clinical domains. It is impossible to represent all of these arrangements within the constraints of the current RBAC system as only legal organisations from NACS are represented in the organisation codes. It is proposed that a number of "virtual organisations" are created that can be used within the URP as the key to access the data appropriate to each shared service. These should legally be a subsidiary of the lead organisation to ensure there is appropriate accountability for their use. Shared Services are discussed in detail in Section 6.

- **Allowing access to SUS for external agencies to SUS by means of RBAC.** The broad principles for this type of access have been outlined in a previous paper to the Board[3].

3.2.7 There are a number of policies and procedures in CfH that have been formulated ahead of the implementation of SUS relating to access to data. Although these are suitable to a large number of scenarios for SUS, the complexities of the organisational working arrangements have tested their applicability and practicality. It is recommended that the SUS Team continue its dialogue with the CfH Information Governance Team to ensure that they are informed of the business processes that SUS is required to support.

3.2.8 As briefly discussed above, there are several external agencies that will require access to SUS data. The practicalities of providing such access need to be investigated, including the provision of smart cards and providing N3 access. Governance rules will be required for the provision of physical access and acceptable use

3.3 Non-technical considerations

3.3.1 As discussed above, any rules controlling the technical governance system must be augmented by organisational controls and rules for appropriate data access and use. It is believed that currently NHS users are notified of their responsibilities regarding access of to and use of data as part of their contracts of employment. The terms of these rules should be confirmed to ensure that they are appropriate for SUS data. It may be necessary to expand on the principles in the Care Record Guarantee and *Confidentiality* to include additional guidance that is specific to SUS. For example:

- Guidance about the mixing of clear and pseudonymised data.
- Guidance about the appropriate use of clear data where it is provided, both during the transitional period and in the longer-term.
- Examples of prohibited data sharing and use.

3.3.2 Where non-NHS parties are allowed access to SUS data there need to be appropriate contractual arrangements that limit the use of that data to those in line with the SUS Information Governance Rules.

3.3.3 Appropriate use of data by both NHS and third parties should be monitored, and there should be a tariff of proportionate sanctions available for misuse.

[3] Access to the Secondary Uses Service; SUS Project Board 15th December 2005

4 Putting the Information Governance Policies and Procedures into operation

4.1 Short to medium-term

4.1.1 In the short-term there are a number of Information Governance Issues that must be settled in order to proceed with the imminent releases. The most urgent of these are the issues surrounding Shared Services, and the transitional and longer-term arrangements to allow access to clear data. These are described in Section 6 with detailed recommendations for the SUS Board.

4.2 Medium to long-term

4.2.1 It is proposed that an appropriate body is formed that can take overall responsibility for SUS Information Governance working within the overall CfH or Care Record Development Board (CRDB) Information Governance setting. This body would be required to:

- Approve/work with the general approach to SUS Information Governance proposed within this paper.
- Formulate SUS Information Governance policies in more detail.
- Liaise with other IG bodies within CfH and the NHS e.g. the RBAC Governance Board, CfH Digital Policy Team in order to ensure that SUS and more central policies are aligned.
- Liaise with external data suppliers.
- Establish detailed procedures for users to follow if they require access to SUS data.
- Oversee and audit the use of (clear) data within SUS.
- Have external visibility e.g. to Patient Information Advisory Group (PIAG), Security and Confidentiality Advisory Group (SCAG), the Health and Social Care Information Centre (H&SIC), etc.

5 Transitional Period

5.1 Need for Transition

5.1.1 As previously identified to, and approved by, the SUS Board[4], there will need to be a transitional period between ceasing the use of NWCS based services providing patient identifiable[5] ('clear') data and the full introduction of pseudonymisation. This transition is required for a range of reasons including -

- Future provision of SUS facilities – until a significant part of the analysis and linkage facilities to be provided as part of the overall SUS developments are in place, analysis will need to be undertaken locally, hence the requirement for data extracts.
- Future provision of alternative facilities – some uses of clear data (eg waiting times tracking) will be met through facilities yet to be provided by Local Service Providers (LSPs).
- Data quality – data quality needs to be of a suitable standard across a range of datasets to enable analysis and linkage between datasets. Currently those standards are not met universally, but will improve with the Data Quality Initiative and the introduction of new facilities (eg SUS copy of patient demographic service) and systems.
- Business continuity – some NHS business would not be able to function with an abrupt changeover to pseudonymised data because key functions, such as checking records or data about patients between organisations or tracking waiting times, could not be undertaken. Lack of such business continuity could impact directly on patients and service provision, eg failure to track waiting times.

5.1.2 The proposed transitional arrangements will also enable the legitimate uses of secondary use data, as set out in the report of the Pseudonymisation Impact Assessment Study, (eg the need for identification of patients in service provision, such as patients at risk derived from 'frequent fliers' analysis) to proceed.

5.1.3 These proposed transitional arrangements therefore apply to NHS organisations such as PCTs, SHAs and associated Public Health organisations, such as the network of PH Observatories.

5.2 Proposed Approach on Clear Extracts & User Roles

5.2.1 SUS will provide access to a SUS Presentation layer. This in turn will enable access to standard reports containing aggregated data and parameter driven reports generated from pseudonymised data. However, to provide the detailed analysis in supporting NHS business functions in the immediate future, local analyses will need to continue to be

[4] Pseudonymisation Impact Assessment Study approved by SUS Board October 2005

[5] Relevant data items are NHS Number, local patient identifier, date of birth, sex, postcode – names and addresses are not included in these records

undertaken. To facilitate this, extracts of commissioning data sets (CDS) type data files from SUS will be provided.

5.2.2 Some of the linkage and analyses applied to records and data derived from CDS can be undertaken on pseudonymised data and therefore the primary source of data from SUS will be pseudonymised. However at present, not all analyses and linkage can be achieved, so access to clear data extracts is necessary. Such provision will be controlled by restricting access to the extract facilities to 'senior' users who should be limited in number in any organisation, typically no more than two; a minimum of two being necessary for business continuity reasons.

5.2.3 Access by SUS users to patient record level data will be restricted to records relating to the organisations on behalf of which they are operating SUS. This will be achieved by functions in SUS deriving the organisations from the NACS codes contained in the RBAC URPs used when accessing SUS.

5.2.4 In addition, there will be a SUS Protocol (see section 5.3) governing the storage and use of clear and pseudonymised data; this will need to be signed on behalf of the receiving organisation and should be incorporated into the local information governance regime. These arrangements will apply to individual organisations and any shared services arrangements (see section 6).

5.2.5 The above proposals will limit who can extract clear data and will require the appropriate management of that data in local NHS environments.

5.2.6 The access to patient identifiable data should be related to the purpose for which the data is required. This means that for PCTs during the transition period, access is needed to the full range of patient identifiable data items, whilst for public health organisations and SHA's access can be limited to files where only the postcode is in a clear form. There may be occasions when public health observatories need access to more detailed data.

5.2.7 For trusts that provide the data, if they require extracts for checking and comparison purposes (previously not possible with NWCS leading to fears and doubt about consistency of data between organisations), then it should be in clear form (as the original data from the trusts will be in that form) in order that pseudonymisation keys are not compromised

5.2.8 It is proposed that SUS Users be broken into two types within their organisational setting, with the types relating to whether access is to pseudonymised data, as a 'standard' user or patient identifiable data as a 'senior' user. Access to the different data will be controlled through RBAC. These proposals are summarised in Table 1.

Table 1 - Data access rights

Organisation Type	SUS User Type	
	Standard	**Senior**
PCTs for Information /PbR	Aggregate & pseudonymised data	Aggregate & pseudonymised data + full clear data
PHO	Aggregate & pseudonymised data	Aggregate & pseudonymised data + clear post coded data
SHA	Aggregate & pseudonymised data	Aggregate & pseudonymised data + clear post coded data

5.2.9 There will need to be consistency of the application of pseudonymisation keys within groups of organisations or users. For example, the CDS data provided to a PCT will need to be consistently pseudonymised so that records relating to specific individual patients utilising NHS services can be linked (eg determining re-admission rates). The pseudonymisation keys across organisations should be different to prevent 'triangulation' of records and hence identification of individuals.

5.2.10 There are potential implications for local NHS systems that are used for the management and analysis of patient record level data. Currently, these systems receive data from NWCS with clear patient identifiers. It is to be expected that these systems will continue to be populated with clear data extracts from SUS in order to provide consistency of the record base and to enable records to be linked eg to determine re-admission rates over different years. It is vital that clear and pseudonymised data are not mixed, as this could compromise the pseudonymisation keys and render the pseudonymisation ineffective. Therefore, parallel databases of pseudonymised records will need to be set up by local NHS organisations to enable routine and basic data analyses to be undertaken.

5.3 SUS Protocol

5.3.1 A SUS Protocol is required to indicate to users of SUS how the outputs of SUS should be used, similar in purpose to the HES Protocol, from which lessons have been gained. The transition outlined above will see the routine extraction of patient identifiable data and the SUS Protocol will need to cater for this. In the longer term, access to and extraction of clear data will be limited and the SUS Protocol will need to be revised to reflect the changing circumstances.

Aim of SUS Protocol

5.3.2 The aim of the SUS protocol is to provide a means of ensuring appropriate handling and use of patient record level data, patient identifiable or pseudonymised, derived or extracted from SUS. Organisations which extract the data from SUS must take responsibility for that extracted data and a mechanism is need to provide assurances about the management and use of extracted data.

5.3.3 Therefore, the SUS Protocol must

- be signed on behalf of any organisation (single or shared) extracting data from or accessing data in SUS and returned to the SUS Board.

- must be incorporated into local information governance arrangements so that the organisation's staff abide by the Protocol.
- apply to all SUS Users.

5.3.4 NHS organisations have extensive experience of handling patient record and sensitive data, together with formal information governance arrangements and should not have problems in conforming to the SUS Protocol. There are processes in place across the NHS for assessing and auditing the operation of information governance in individual organisations.

Outline SUS Protocol

5.3.5 An outline SUS Protocol is set out in Table 2 to provide the basis for the development of a detailed Protocol.

Table 2 - Outline SUS Protocol

Protocol Element	Detail
Purpose	The SUS Protocol is a component of the SUS Information Governance arrangements, concerned with protecting the confidentiality of patients. The protocol sets out guidelines for SUS Users on handling and management of data and the release of data, particularly during the transition to the wider implementation of pseudonymisation. The Protocol is also a visible symbol indicating the appropriate handling of patient confidential data.
Status	This Protocol forms part of the Information Governance arrangements of any organisation extracting data from SUS and is binding on all SUS users.
Audience	All organisations using SUS and all individual users of SUS
SUS Data & Data Extracts	The SUS warehouse holds data about individual patients and their care in a CfH standard secure environment. The data is held in clear (patient identifiable) and pseudonymised forms. It is possible to extract clear data from SUS for legitimate purposes supporting medical and healthcare delivery purposes and, during a transitional period, for NHS business continuity. Access to these extracts is controlled by users being registered for a suitable business function in RBAC. Unauthorised access must not be allowed by enabling other staff to share legitimate SUS access.
Data Storage	Data extracted from SUS must be stored in a secure environment with controlled and restricted access. Patient identifiable and pseudonymised data extracts must be stored separately.

Protocol Element	Detail
Data Management & Usage	The management and use of patient identifiable and pseudonymised patient record level data must be undertaken in such a way the different data types cannot be mixed and cause the pseudonymised records to be compromised. Similarly, pseudonymised records must not be analysed in such a way as to deduce the identity of individual patients.
Data Release	Patient identifiable data can only be released within the NHS environment for the express purpose of supporting, directly or indirectly, the provision or delivery of health care services to patients. Guidance on handling SUS data, eg small numbers, must be adhered to.
Audit	The management of patient identifiable data extracted from SUS (in terms of where stored, who released to, etc) must be capable of being audited in order to ensure the effective implementation of this protocol in the local organisation.

6 Shared Services

6.1 Introduction

Formatted: Bullets and Numbering

6.1.1 There are specific information governance issues in relation to information and commissioning services operated on shared basis between and across NHS organisations and actions need to be taken to enable use of SUS by 'shared services'. Resolution of the issues raised by shared services is urgently required to enable the timetable for decommissioning the existing NWCS data distribution service to be met.

6.2 Context of SUS & Shared Services

Formatted: Bullets and Numbering

6.2.1 SUS will provide access for authorised users to data relating to patients and service users in the form of CDS for patient activity, (eg admitted patient care, out patients, waiting lists, A&E etc) extracted and derived from Trusts' operational and clinical systems (eg PAS), and in future from the NHS Care Records Service.

6.2.2 As indicated in Section 5 on Transition, in the short term, relevant data in SUS will be extracted and analysed on a local basis whilst in the longer term, analysis will take place within SUS itself. The data so extracted from SUS will be used for a variety of purposes, such as supporting the day-to-day business of the NHS in terms of commissioning services, performance management of services, audit, public health, etc. The immediate users of SUS will be members of NHS PCTs, Trusts, SHAs and associated public health organisations and will typically be members of information, finance and public health.

6.2.3 Some NHS organisations have formal or informal arrangements to share the necessary information management and analysis services or commissioning services in order to achieve economies of scale through, for example, reductions in duplication of effort, creating critical mass of scarce skills, sharing overheads, reducing management costs etc. With these shared services, staff in a 'lead' organisation or a team/unit attached to that organisation, undertake work on behalf of those other organisations. Examples of such shared services are Health Informatics Services, based in a PCT and serving a group, of say, 3 or 4 of PCTs or based in a SHA based providing information management and analysis services to all the SHA's PCTs; or specialist commissioning teams, for example commissioning cancer services across a group of 15 PCTs.

6.3 Accessing Data from NWCS and SUS

Formatted: Bullets and Numbering

6.3.1 Currently, shared services will have access to data provided from the participating organisations and from NWCS; for PCTs the NWCS data will be patient identifiable or clear data whilst for an SHA, patient data would be pseudonymised. Data from NWCS is provided on a 'push' basis that is after data is processed by NWCS, output is produced on a PCT basis and resulting output files are sent to relevant PCTs according to the contents of an Address Grid held within the NWCS system.

6.3.2 When SUS replaces NWCS as the provider of CDS data for PCTs, the means of accessing the data will change. Users in PCTs or shared services will have to 'pull' the data from SUS. In order to do this, users will have to log on to SUS through RBAC facilities using their unique user identity and their specific role profiles. For users extracting data relating to their own organisation, this should be straightforward. RBAC utilises the organisation codes derived from the National Administrative Codes Service (NACS), which identifies all formal and legal NHS bodies and others, which have been registered with the services.

6.3.3 Similarly, access facilities need creating for SUS users working on behalf of shared services to enable them to access relevant CDS and only those CDS. Whilst this is more complex in terms of the applications managing the data, the requirements for this have been clarified and application modifications are being developed. To support shared services, such organisational arrangements will need to be registered with NACS in order to enable recognition of users through RBAC and to enable the applications to allow access to appropriate data files.

6.4 Identifying & Registering Shared Services

Formatted: Bullets and Numbering

6.4.1 To access relevant data in SUS, identification of Shared Services is required at two points

- In RBAC – to allow legitimate access into SUS for the Shared Service together with identification of the relevant business functions associated with the user.
- In SUS – to enable the relevant data, marts and files associated with the organisations served by the Shared Service to be presented for use by the end user.

6.4.2 For RBAC, a NACS code is required for each Shared Service, which means that each Shared Service must be registered with NACS.

6.4.3 For SUS, there is

- A minimum need to access and maintain relationships between organisations and shared services to support NWCS Decommissioning;
- A further need to extend this to contain contextual information about which services (eg commissioning cancer), and therefore types of data, are relevant to the relationships.

6.4.4 For SUS, in the short-term, the CRISP component of the NWCS Address Grid facility (which is concerned with information which organisations should receive output datasets, known as copy recipients) *may* provide the basis for identifying organisational level relationships between the Shared Services and the organisations it supports. The NWCS Address Grid has been replicated in terms of functionality as a stand-alone application for SUS and can be utilised by applications within SUS. It is expected that the resulting SUS Address Grid will include the vast majority of such arrangements operating in the NHS and this can act as a starting point for enabling access to data (as it the basis on which data is currently distributed). This could provide the facilities to meet the minimum need outlined above in the third bullet point.

6.4.5 Beyond the immediate requirements, consideration needs to be given to how best to register and maintain records of the complex inter-organisational arrangements that constitute shared services for use with SUS and, if possible, how links can be established to the systems and records used by NACS to act as an information service to SUS about shared services.

6.5 Information Governance

Formatted: Bullets and Numbering

6.5.1 Suitable information governance mechanisms need to be in place to support the use of SUS by shared services. It is proposed that criteria be established for inter-organisational arrangements to be considered bona-fide shared services in terms of legitimate right to access CDS type data. These criteria would be based on meeting relevant Information Governance policy aspects set out in this paper and would include

- the 'legal' basis for the Shared Service (eg contractual relationship with which organisations, identifying which is the lead organisation, suitable corporate governance arrangements, such as a Board). This would need to apply to all types of Shared Service (eg HIS, Commissioning, SHA, Independent Sector & Joint Ventures).
- evidence of locally auditable information governance policies to ensure controlled access to identifiable data in a locally secure environment. This could be achieved through the IG Checklist (a self assessment mechanism), which in turn links directly to the Healthcare Commission's assessment of organisations to give independent verification of achieving suitable information governance standards.

6.5.2 In addition, it is imperative that the credentials of the shared services already registered with NACS are checked to ensure that the relationships are valid prior to the commencement of access to SUS for data extraction.

6.6 Supporting Actions and Timetable

Formatted: Bullets and Numbering

6.6.1 The actions required to enable Shared Services to access SUS include -

- Notification of all known and potential shared services of steps to take associated with the change from NWCS to SUS including
  - Identification of lead organisation
  - Checking against SUS Shared Services organisation criteria
  - Registration by the lead organisation of the shared service with NACS
- Liaison with NACS about new registrations
- Development of functionality within SUS to handle shared services
- Continued liaison with CfH RBAC Team.

6.6.2 The timetable required for these actions is as below

- Notify NHS organisations about the approach on shared services - end of February
- Registration of shared services with NACS – end of March
- SUS shared services functionality operational – June 2006.