Data Management Strategy

Data Management Strategy
(v1.0, February 2015)

Scope

1. This document focuses primarily on the internal data management objectives of the CCG over the next three years. Due to the evolving nature of legislation in this area, the strategy and progress against objectives outlined here will be reviewed annually.

2. Future iterations of this document will expand the scope to include broader considerations, such as the role CCGs must play in ensuring improvement in provider data quality, completeness and integrity.

Executive Summary

3. This document describes the primary strategic objectives for CCG-led data management in Nottinghamshire (excluding Bassetlaw) for the next three years. A high-level outline of the tactics proposed to deliver these objectives is also presented.

4. With increasingly challenging financial environments and cost saving plans affecting the NHS and Local Authorities, information is critical to helping primary care deliver the necessary improvements with limited resources available.

5. The vision reflects the CCGs’ strategic objectives, driven by the quality of care, patient health and care outcomes, the reduction of inequalities and increases in productivity and efficiency.

6. Our key strategic objectives are to:

   a) Ensure that information is held in the most appropriate environment for clinical decision-making (e.g. risk of admission data held principally in clinical information systems such as EMISWeb and SystmOne).

   b) Enable Practices to easily extract from and import to their clinical systems, to support data integration and processing.

   c) Integrate information from primary care, community care and acute providers to provide a global view for direct patient care, commissioning and performance monitoring.

   d) Provide a secure technical- and governance framework to facilitate cross-organisational sharing of data for direct care as well as for secondary uses.

   e) Drive improvement in data quality.
Data Management Strategy

7. To accomplish this we will:

   a) Use innovative techniques to collect and integrate data from our different clinical (and, where possible, social) sources.

   b) Guarantee the security of our data and the rights of our patients.

   c) Support decision-makers by using valid statistical techniques, relevant indicators and clear information displays.

   d) Develop data management models which are flexible by design, to mitigate against externally-forced change requirement, to the extent possible.

8. Our approach to collecting data will be to:

   a) Encourage the use of structured data entry in primary care through using cross-Nottinghamshire shared templates and/or codes.

   b) Establish a semi-automated local extraction and upload process for primary care clinical information systems such as EMISWeb and SystmOne.

   c) Semi-automate the process of getting data from community (such as CHP, CityCare, PICS and Red Cross) by providing bespoke tools to pre-process, validate and pseudonymise data at source, prior to upload.

   d) Work with our major acute providers to get early data following admissions and attendances.

   e) Integrate data from the different CCG projects (e.g. EPaCCS, bowel cancer screening, referrals) with that from our primary sources to produce a richer picture of patient care for our different needs.

9. To guarantee the security of data and rights of our patients we will:

   a) Use pseudonymisation techniques to safely transfer, integrate and store data across the local health and social care community.

   b) Use Role Based Access to ensure that end-users have the right level of access to personal confidential data, pseudonymised, anonymised or aggregated data to fit their rights and needs.

   c) Use an explicit permissions framework visible to end-users showing which roles and individuals have access to each information display.
To support decision-makers in practices, CCGs and the wider community we will:

- Put into place electronic Data Sharing and Processing Contracts to enable us to respond quickly to changing requirements, with clarity.
- Use smartcards to bring closer the ideal of single sign-on and circumvent the barrier that login provides to accessing our data.
- Use a modern, supported data platform, content management system and programming environment to deliver our products.

10. To support decision-makers in practices, CCGs and the wider community we will:

- Analyse data using statistical techniques to balance out differences in age, sex and deprivation.
- Display data with confidence intervals so users can distinguish real differences from noise.
- Show comparative data from across our community (and where possible from beyond it) so we can understand the relative position of each practice, locality, CCG, ward and district.
- Use populations that closely represent reality at the time of each analysis.
- Automate graphing and mapping of data for each indicator we analyse.
- Allow clinicians to drill down from aggregate findings into the specifics of individual cases to support their need to change processes and audit their care.

**Enterprise Data Extraction for Primary Care**

11. With support from Connected Nottinghamshire, a funding bid is being developed to set up a semi-automated data extraction service for EMIS Web and TPP SystmOne Practices. The aims are:

- To provide a mechanism for routinely extracting key data from primary care systems.
b) To make this process secure and semi- or fully-automated.

c) To enable processed data (e.g. risk of admissions scores) to be loaded back into the clinical systems.

d) To support integration between EMIS and TPP Practices.

e) To promote the use of standardised extract specifications and templates.

12. The purposes and benefits of doing this are aimed squarely at direct patient care. For example, risk of admissions data would best be held within a clinical system. However, in order to generate the scores, data must first be extracted, pseudonymised and then processed through a risk-profiling algorithm within a secure environment. During a multi-disciplinary team meeting, the eHealthScope Risk of Admission Register offers a much more efficient environment for rapid reviewing and updating of those records but, thereafter, the results should be loaded back into the clinical system.

13. For this strategy to succeed, it will require the trust and support of GPs, underpinned by robust Data Sharing and Processing Contracts.

14. To support evolving requirements, these Contracts will be developed in such a way that amendments may be managed electronically via an interface in eHealthScope. The benefits of such an approach are:

a) Changes can be implemented quickly with minimal manual administration.

b) Additions and amendments can be highlighted for clarity, leaving other details in the contracts unchanged.

c) The full, agreed contract may be compiled and distributed electronically.

d) The version currently in force is always available for inspection online.

e) Previous versions may be captured to show a ‘timeline’.

f) There is future potential to link consent status in the contract directly to technical processes.
**Controlled Data Flows and Pseudonymisation**

15. There are three key s251 exemptions actively in force which underpin the utilisation of Patient Confidential Data (PCD) by CCGs:

<table>
<thead>
<tr>
<th>Scope</th>
<th>Reference</th>
<th>Expiry Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Stratification</td>
<td>CAG 7-04(a)/2013</td>
<td>30 Apr 2015</td>
</tr>
<tr>
<td>SUS and other commissioning data from HSCIC (ASH)</td>
<td>CAG 2-03(a)/2013</td>
<td>30 Apr 2015</td>
</tr>
<tr>
<td>Invoice Validation (CEFF)</td>
<td>CAG 7-07(a)(b)(c)/2013</td>
<td>30 Apr 2015</td>
</tr>
</tbody>
</table>

16. Although the s251 support enables NHS numbers to be processed within an ASH without further pseudonymisation, there is clear intention that such arrangements should be regarded as temporary.

17. As a national solution which might remove the underlying need to provide this s251 support to CCGs is not yet in place, NHS England has secured the extension of the current s251 exemptions until at least April 2015.

18. A public consultation on the future of NHS data flows and ASH arrangements was recently undertaken by NHS England¹. Latest indications suggest that the forthcoming General Election is likely to delay the announcement and implementation of any changes in legislation. A robust local strategy must aim to predict the most likely outcome whilst mitigating against significant foreseeable risks.

19. The Data Management Team (DMT) continues to actively pursue implementation of the ‘Nottinghamshire Model’ within the DSCROs – a framework in which data are retained in the local data warehouse but pseudonymisation and re-identification functions are managed within the secure and legal environment of the HSCIC. Under existing rules, this would allow the CCGs to operate without specific s251 support and without the burden of maintaining ASH status.

20. Following a dedicated site visit on 1 August 2014 from Ming Tang, Director of Data and Information Management Systems, our approach has been re-endorsed by NHS England and continues to attract commitment from the DSCROs. Maintaining this close relationship with the emerging DSCRO infrastructure is strategically important to ensure that CCGs are both protected from unfavourable policy changes whilst also in a position to benefit from positive national developments.

21. Due to overwhelming demand on DSCRO resources, the pace of fully implementing the trialled solution has been slow. To mitigate this, in anticipation that future legislation will likely permit CCGs to continue to receive NHS numbers, and as best practice, local pseudonymisation capability has been implemented across the data warehouse and

Data Management Strategy

eHealthScope. This was first introduced for Risk Stratification, where the real NHS number is only re-identified for direct clinical purposes, and has recently been extended across all functions within eHealthScope.

22. The ‘Caldicott 2’ review\(^2\) highlighted the real importance of being able to facilitate safe, secure and legal flows of data between direct care providers (for example, Community Pathway providers and GPs). This was underscored again in a communication from the Department of Health in November 2014\(^3\) calling for wider sharing for direct care, where good IG is seen as the enabler rather than as a barrier.

23. In the First Year Report on the Caldicott 2 review, published by the Independent IG Oversight Panel in December 2014\(^4\), Dame Fiona Caldicott expressed regret that the recommendation for “cultural change” towards information sharing, which was made in the wake of the Francis report on Mid Staffordshire NHS Foundation Trust, had “only emerged in parts of the system”.

24. In some cases, support for sharing is currently in place via s251, but these exemptions cover data flowing via the HSCIC only. This creates significant complexity and capacity bottleneck which, in reality, renders much of this work impracticable and too slow for CCGs to tackle effectively in this way alone.

25. To alleviate this, the DMT has developed additional methodologies based upon ‘pseudonymisation at source’. This approach is currently the frontrunner among solutions being considered at a national level with the aim of moving the NHS away from reliance on s251 support for data processing.

26. The key principles of pseudonymisation at source are that:

a) Data does not leave the Data Controller in its clear form.

b) The Data Controller does not require a specific legal basis to process ‘their’ data for the purpose of de-identification.

c) Although control measures and restrictions still apply, data may be shared and processed more widely once in pseudonymised form.


d) Data may be re-identified to relevant care professionals where there is a legitimate clinical relationship and justification for doing so.

27. The fundamental hitch which prevents more widespread use of pseudonymisation at source is that, in almost all examples of medical computing, datasets are not used in isolation but must be joined with others. Typically, joins are made using the NHS number; however, this is not possible with pseudonymised data unless the datasets to be joined have been pseudonymised in precisely the same way.

28. If all data processors shared a common ‘key’ for pseudonymising and re-identifying data, such joins of shared data would be possible, but this would defeat the purpose of de-identifying the data in the first place.

29. To circumvent this problem, the DMT has developed a system- (or ‘black box’) process which translates between pseudonyms chosen by the data provider and those which are used throughout the data warehouse. This is achieved in such a way that no PCD is seen or touched by the DMT, and our pseudonymisation keys are not shared with the data provider.

30. Where a specific legal basis to re-identify data can be verified (e.g. a GP accessing the NHS number of a patient registered with them and who has returned a high risk of admission score), this can be facilitated and audited via eHealthScope.

31. The Information Commissioner’s Office (ICO) has recently issued a public response to queries raised by the Department of Health pertaining to pseudonymisation standards. The ICO supports the view that a case-by-case consideration is required in order to determine whether or not pseudonymised data constitutes “(sensitive) personal data”, in the context of the Data Protection Act, 1998. This will depend upon the full content of the record and the availability of other sources, and therefore whether it is “reasonably likely” that individuals could be identified in specific circumstances.

32. Positive discussions about applications of this new technique are currently ongoing with IG leads across various local organisations and it is hoped that practical trials can begin in February 2015.

33. An added benefit of our approach to pseudonymisation at source is the ability to run data quality checks before data leaves the provider. This will help CCGs to achieve their commitments towards improving data quality, as the best time to identify and correct any such issues is whilst the data is in the hands of the provider. We may apply any level of

---

rigour deemed necessary, from automatic corrections for straightforward issues, to issuing warnings about detected issues, or even preventing a dataset from being uploaded to us at all until a particular quality threshold is met. Requirements will need to be set and reviewed on a case-by-case basis, dependent on practical considerations.
34. The creation of the shared Data Management Team and evolving legal basis of Health and Social Care data flows necessitates that the respective roles of NHIS and the DMT are better clarified. This is not entirely straightforward, as administrative boundaries are currently tied to legacy systems which are in the process of migration: not least, to address this very issue.

35. Whilst it is essential that the DMT finds ways to work much more closely with NHIS at a strategic level, it is equally important that NHIS maintains a distinct identity in order to provide third-party Data Processor services to a range of customers, including our CCGs and GP Practices.

36. This distinction may be particularly pertinent for third-party pseudonymisation techniques, discussed above, since the party holding the keys may be deemed to have a higher “reasonable likelihood” of re-identifying personal data, which would otherwise be considered outside the scope of the Data Protection Act in its pseudonymous form. For this reason, we do not currently propose the TUPE of staff from the NHIS BI Support Unit into the DMT.

37. There is widespread confusion, nationally, about the role of CCGs as commissioners (akin to former PCTs), and as a consortium of GP Practices. Practices rely on quality, timely and accessible information upon which to base clinical decisions and meet their obligations, and CCGs play a vital part in supporting this. Often, these requirements are quite different to what is traditionally considered to be commissioning activity.

38. As statutory bodies, CCGs are able to enter into legal contracts and therefore to act as Data Processors for other organisations, including member GP Practices. This may involve the legal processing of data on behalf of Practices which the CCG does not have the right to hold in its own capacity as a commissioning organisation (i.e. within the CCG Accredited Safe Haven). To avoid ambiguity about the purpose of data flows and underpinning legal basis, respective Data Processing contracts will be continuously reviewed and revised. The ability to utilise another organisation (e.g. NHIS or the DSCRO) as a mutual intermediary Data Processor, as described by the “Nottinghamshire Model”\(^6\), is a key risk mitigation element within our strategy.

39. All projects undertaken by the Data Management Team shall include a review of the IG framework, with those details recorded within eHealthScope alongside the content. This will enable clear and straightforward auditing of the identified Data Controller(s), Data Processor, purpose and access permissions.

---

\(^6\) “An Integrated Data Model for Health & Social Care” – Carl Davis & Dr Mike O’Neil, 2014
## Risk Assessment

![Risk Assessment Table]

<table>
<thead>
<tr>
<th>No.</th>
<th>Identified Risk</th>
<th>Likelihood</th>
<th>Impact</th>
<th>Score</th>
<th>Controls / Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SQL Server 2008 Standard Edition insufficient for the technical challenges</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>SQL 2012 Enterprise option being costed by NHIS, possible opportunities with transformation funding bids</td>
</tr>
<tr>
<td>2</td>
<td>DSCRO unable to provide data</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>Maintain relationship with other DSCROs / CSU as backup</td>
</tr>
<tr>
<td>3</td>
<td>DSCRO unable to implement Nottinghamshire Model satisfactorily</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>Move more of the pseudonymisation process into the DSCRO environment, with probable additional cost.</td>
</tr>
<tr>
<td>4</td>
<td>Pseudonymisation forced to take place at HSCIC and technical difficulty in implementing this at the speed necessary</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>Early trials show promise.</td>
</tr>
<tr>
<td>5</td>
<td>Unforeseen changes in IG requirements</td>
<td>4</td>
<td>3</td>
<td>12</td>
<td>Implementing best practice of pseud-at-source and appropriate data processing contracts is a step ahead of current requirements.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Permissions framework in eHealthScope ties user permissions (assigned by the Data Controller) to specific functionality within each module. Also tied explicitly to the data that can be returned and it’s type (anon, pseud, weak pseud or clear). Will be able to build in expiry dates with reminders for DCs to review permissions they have assigned.</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Data is disclosed to people with insufficient permissions</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Practices uncomfortable with data extraction from their CIS</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Financial risk if CCGs withdraw funding for DMT</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Clinical involvement and gaining trust, following a long tradition in eHS. Guarantees of restricted access, transparency and emphasis on direct patient care.

Many examples of money being saved through use of information (Cardiology Outpatient project, Risk of admission register, Reductions in referrals with release of referrals log)