Project / Programme	**Data Dissemination Services**	Project	**Data Sharing Audits**
		Status	**Approved**
Director	**Terry Hill**	Version	**1.0**
Owner	**Rob Shaw**	Version issue date	**15/08/2016**

NHS Digital Audit of Data Sharing Activities: Health IQ

Contents

Executive Summary

This document records the key findings of a technical audit of Health IQ following NHS Digital concluding the company had breached its Data Sharing Agreement by holding and processing data outside of the UK. As a result of this breach Health IQ had been required to delete NHS Digital data from the non-UK systems and to cease processing data in its London office. As a result of the incident, Health IQ has moved some of its infrastructure to a UK data centre.

The purpose of this technical audit was to examine the new storage and processing arrangements implemented by the company to address the cause of the breach prior to any decision being taken by NHS Digital as to whether the company could start to receive data again. However, as the organisation had not been previously audited by NHS Digital the Audit Team extended the scope to include other areas normally examined as part of a data sharing audit.

Due to the nature of the audit, the process by which it was conducted and recorded differs from the published data sharing audits conducted by NHS Digital since August 2014. Specifically, this audit was conducted over an extended time-frame which allowed Health IQ to address the issue regarding small number suppression in its on-line tool. As a result, conclusions are presented as statements rather than as non-conformities or observations based on the evidence presented during the on-site visit.

The Audit Team has concluded from this extended review that the storage and processing arrangements being established are reasonable for the nature of the data and have addressed the cause of the identified breach. However, as no data has yet been released to the company and therefore supporting evidence is lacking in some areas, a further audit is recommended. This further audit would also look at how the organisation is handling new data.

In summary, it is the Audit Team’s opinion that based on evidence presented during the audit there is low risk of inappropriate exposure and / or access to data provided to Health IQ by NHS Digital.

1 About this Document

1.1 Introduction

The Review of Data Releases by the NHS Information Centre produced by NHS Digital Non-Executive Director Sir Nick Partridge recommended that NHS Digital should implement a robust audit function that will enable ongoing scrutiny of how data is being used, stored and deleted by those receiving it.

In August 2014, NHS Digital commenced a programme of external audits with organisations with which it holds data sharing agreements.

1.2 Background

In May 2016 Health IQ was found to have breached its Data Sharing Agreement with the storage and processing of NHS Digital Hospital Episode Statistics (HES) data taking place within the Republic of Ireland. As a result of this breach the company was instructed by NHS Digital to delete data from the Ireland servers and to cease processing data in its London office.

The NHS Digital investigation also recommended that a detailed technical audit be carried out of Health IQ's new processing and storage arrangements before being permitted to receive fresh data under the existing agreement.

1.3 Purpose

This report provides an evaluation of how Health IQ has adapted its storage and processing arrangements in order to conform to the requirements of a Data Sharing Framework Contract and a Data Sharing Agreement.

This report provides a summary of the key findings.

1.4 Presentation of Findings

As this technical audit involved a more detailed and protracted review of the new storage and processing arrangements and Health IQ's implementation of small number suppression, this report does not use the established approach taken to the reporting of data sharing audit findings (i.e. nonconformities and observations). Sections 2 and 3 do, however, present findings both from the on-site visit and at the time of writing in order to present a balanced view.

1.5 Audience

This document has been written for the NHS Digital Director of Data Dissemination Services. The report will be published in a public forum.

1.6 Scope

The audit considered the fitness for purpose of the new storage and processing arrangements being implemented by Health IQ and whether they posed any risk to patient confidentiality or to NHS Digital.

1.7 Audit Team

The Audit Team was comprised of senior certified and experienced ISO 9001:2008 (Quality management systems) and ISO 27001:2013 (Information security management systems) auditors.

2 Audit Findings

This section presents the key findings arising from the audit, which included:

- an on-site visit to Health IQ's offices in London on 13th and 14th July; and
- a review of Health IQ's on-line tool (Vantage) to ensure that the level of suppression and re-identification were in line with the Hospital Episode Statistics (HES) Analysis Guide[1].

As no data has been cleared for processing and storage by NHS Digital since the incident, the new practices employed by Health IQ could not be fully audited. This restriction also acknowledges that implementation changes are in the process of being finalised.

2.1 Processing and Storage of Data

HES data is downloaded from NHS Digital via Secure Electronic File Transfer (SEFT) and saved onto an external USB encrypted drive. The downloaded data is then transferred to an EFL (Extract-Transform-Load) encrypted server and deleted from the encrypted drive. Processed data is saved in a local data warehouse for testing by Health IQ analysts. All processing of the raw HES files is done on this local server. A copy of the processed HES data is saved to an external encrypted backup drive.

An extract of the validated data is then copied (via an encrypted channel) to a SQL database held in a UK datacentre. Only when this transfer takes place are the firewall ports opened; normally the local server is fully offline. The server in the datacentre houses both the tool (Vantage) and the processed data. The data centre server also has a dedicated back-up server.

As the selection of the new data centre is recent not all of the processing/testing/working practices have been finalised. The ability to process new data will engage some of the activities covered in the above paragraphs for which there is little or no evidence.

Health IQ analysts use the processed data to generate reports for clients, though only aggregated data will be presented. Health IQ analysts use individual encrypted laptops to generate these reports / Excel tables.

A number of vulnerability tests have already been conducted on the platform in the new datacentre and the severe findings rectified. Further penetration testing is to be conducted once Health IQ has received a full data set.

Valid customers are able to access the on-line tool to display aggregated data. As part of this access the company has a standard sub-licensing agreement. The Audit Team noted that the sub-licensing agreement had no date of signatory recorded therefore making it difficult to have unique identification. A scanned signature is also used.

Usage reports are produced according to the customer and the company reported that annual customer meetings are held.

[1] www.hscic.gov.uk/media/1592/HES-analysis-guide/pdf/HES_Analysis_Guide_Jan_2014.pdf

2.2 Access Controls

To provide business continuity the NHS Digital SEFT login credentials have been shared between two members of staff. The company was informed NHS Digital could not condone such an activity and the main person should request new login details for personal use. If this person is unavailable at the time of downloading data, Health IQ should contact the NHS Digital SEFT team for assistance.

Access to the servers is controlled with only two members of staff currently having access. No analysts have access at present due to the restriction on processing data imposed by NHS Digital.

The Audit Team reviewed the arrangements by which nominated internal staff have physical access to portable USB encrypted drives.

2.3 Data Destruction

A 3rd party company is used to destroy redundant equipment. However, the process around equipment destruction needs to be tightened and the various pieces of information reconciled. For example, a number of hard discs were recently destroyed on site by the 3rd party company. Supporting this destruction were 11 completed internal destruction forms, 12 items were identified in company's asset log, 12 items were detailed by the 3rd party supporting the destruction certificate, and the number 13 was stated on the collection notice.

The company also needs to consider what to do in the event of a hard disc failing within a laptop that is under warranty. That is, is the disc retained or is it returned?

2.4 Operational Management and Control

The company has a number of procedural documents, though many of these are relatively new and are still at draft. Furthermore some documents reflect how the company wants to act rather than how its acts at the moment. New templates have also been developed. Consequently there is no or little evidence to support some of the newer working practices. For example, until Health IQ starts to process data it will not invoke the process around giving analysts access to the HES data. The wording needs to be checked in some documents as it is potentially misleading and should be modified. Version control also needs to be improved.

Health IQ is to seek certification to ISO 27001:2013 at some point next year and is investing time and effort into developing an Information Security Management System (ISMS) supported by a range of policies, processes and procedures. The Audit Team would expect future documentation to have a correct and consistent approach to document management.

As part of the revision to working practices the approach to risk management is going to change. Some work has already been done on the approach but no implementation of the new approach on risk assessment and risk treatment has been done.

To oversee information governance (IG), a new IG committee has been created. So far it has only had its inauguration meeting, but if run properly and frequently it should ensure appropriate controls are being implemented and monitored. For example, review of folder access.

Health IQ is looking at how to undertaken internal audits and an audit programme has yet to be produced.

All new staff are expected to undertake the NHS Digital IG training. Certificates for some staff were provided to the Audit Team. After this there is less formality around refresher training and there were some important omissions with respect to training for certain members of staff, for example, the Senior Information Responsible Officer (SIRO). The new starter checklist is to be revised to state that the starter will not be able to access data until the IG training has been completed.

The Audit Team noticed that some administration staff were leaving laptops unlocked when leaving the office. This situation was pointed out to the Health IQ manager to address.

The company's data protection policy only requires the Information Commissioner's Office (ICO) to be informed in the event of a data breach. The Audit Team stated that the need to inform NHS Digital as required by the contract / IG Toolkit should be added to the document. The company is maintaining an IG incident log which included the incident initiating this audit.

As the company is running a Unix based server, the Windows laptops cannot be managed centrally. IT Support is responsible for laptop builds and for any changes. The company stated that laptops do not have local administration rights. The free edition of Avant anti-virus is installed on the laptops and the Audit Team found virus definitions to be up-to-date.

The local server is regularly patched and subject to period review. Event logs, firewall logs and SQL logs are retained. The formality and periodicity of these reviews have still to be established.

The company is still looking at what other controls it may need to implement, for example, the use of tools such as Blancco for software wiping.

2.5 On-line Tool

A review of the on-line tool, Vantage, was undertaken. This review found issues with respect to the suppression of small numbers. Namely, whilst any number 1 to 5 is shown as 5, if a query results in say 8 and the query is then widened, for example, "show result over two months", 6 and 5 could be displayed (the 5 representing the hidden 2 which can be calculated from 8 - 6).

A solution to this issue has been proposed by Health IQ which has been accepted by NHS Digital.

3 Conclusions

This section presents the key findings of this audit based on the:

- findings from the on-site visit and tool review; and
- revised findings based on subsequent discussions and reviews.

The main findings are:

- All processing of the raw HES files is done on a server in the company's London office. Processed data to which customers have access through an on-line tool is now held in a UK datacentre. Customers are only able to display aggregated data.
- The approach to small number suppression within the tool required further investigation as it was possible to calculate small numbers in certain situations. A solution has been agreed with NHS Digital.
- A number of vulnerability tests have been conducted on the platform in the new datacentre. Further penetration testing is to be conducted once the company has a full data set.
- The company is to seek certification to ISO 27001:2013 at some point next year and is investing time and effort into developing an ISMS supported by a range of policies, processes and procedures.
- As the move to a new data centre is recent not all of the processing/testing/working practices have been finalised. As a result there is little or no evidence for certain practices.
- The wording is to be checked in some current documents as it is potentially misleading and should be modified. Version control also needs to improve.
- As part of the revision to working practices the approach to risk management is going to change. Some work has already been done on the approach but no implementation of the new approach on risk assessment and risk treatment has been done.
- The company is still looking at what other controls it may need to implement, for example, the use of tools such as Blancco for software wiping.
- A new IG committee has been created which if run properly and frequently should ensure appropriate controls are being implemented and monitored.
- All new staff are expected to undertake the HSCIC IG training; certificates were provided. After this there is less formality around refresher training and there were some important omissions with respect to training for certain members of staff, for example SIRO.
- Reiteration of the need to lock unattended laptops is required as this is not practised consistently by staff.
- The new starter checklist is to be revised to state that the starter will not be able to access data until the IG training has been completed.

- To provide business continuity the HSCIC SEFT login credentials have been shared between two members of staff. The company was informed HSCIC could not condone such an activity and the main person should request new login details for personal use.
- The process around equipment destruction needs to be tightened and the various pieces of information reconciled. The company also needs to consider what to do if a hard disc fails within a laptop that is under warranty.
- On the company's sub-licensing agreement no date of signatory is recorded therefore making it difficult to have unique identification. A scanned signature is also used.
- The company's data protection policy only requires the ICO to be informed in the event of a data breach. The Audit Team stated that the need to inform the HSCIC as required by the contract / IG Toolkit should be added to the document.

As the storage and processing arrangements were relatively new, there was a lack of evidence for some of the processes. It is proposed that a follow-up audit is conducted in a few months to review the final system configuration and to review evidence arising from the processes described by Health IQ and the records being generated by the system, for example access logs.

Notwithstanding the above statement, the Audit Team is content with the new storage and processing activities defined by Health IQ during the on-site visit.