Background Quality Report:

Diagnostic Imaging Dataset

We are the trusted national provider of high-quality information, data and IT systems for health and social care.

www.hscic.gov.uk
enquiries@hscic.gov.uk
@hscic

Author:	**Secondary Care**
Responsible statistician:	**Nicola Dawes, Section Head**
Version:	**V1.0**
Date of publication:	**June 2015**

Contents

Version History

Version	Date issued	Summary of changes
1.0	June 2015	Initial version

Please contact us via enquiries@hscic.gov.uk for earlier versions of the Background Quality Report

Introduction

This document constitutes a data quality report for the Diagnostic Imaging Data set (DID). The DID is a central collection of detailed information about diagnostic imaging tests carried out on NHS patients, extracted from local radiology information systems and submitted monthly. DID data can currently be accessed through the following routes:

- A monthly publication of DID data produced by NHS England. This can be accessed here: http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/

- A monthly publication of DID data linkage to HES data. More information on HES-DID linkage can be found here: http://www.hscic.gov.uk/hesdid

- Aggregate views of DID data through iView, available to employees of NHS and NHS-affiliated organisations. To request access to iView for DID, please email enquiries@hscic.gov.uk .

- Data extracts available through the Data Access Request Service[1]. The data provided by the service is normally pseudonymised and identifiable data is only provided when there is a lawful basis to do so.

The aim of this document is to provide an overview of the data quality issues relevant to the dataset itself, rather than any specific publication or presentation of the data.

Context

Background to the Diagnostic Imaging Dataset

The DID captures information about referral source and patient type, details of the test (type of test and body site), demographic information such as GP registered practice, patient postcode, ethnicity, gender and date of birth, plus items about waiting times for each diagnostic imaging event, from time of test request through to time of reporting.

All NHS providers are required to submit DID data to the HSCIC as mandated by ROCR, the predecessor of the current Burden Advice and Assessment Service (BAAS)[2], and Monitor[3]. The HSCIC collects the DID on behalf of NHS England[4] who provide further information about the data set on the NHS England DID web page[5].

Those who submit DID data upload it securely to the DID system via a dedicated DID website[6]. The site provides the submission timetable and further information and guidance on the data set.

The DID enables analysis of demographic and geographic variation in access to different test types and different providers. It also provides useful information for policy makers, commissioners and researchers. Further information on the benefits of DID are available via the DID website.

---

[1] http://www.hscic.gov.uk/dars
[2] http://www.hscic.gov.uk/baas
[3] https://www.gov.uk/government/organisations/monitor
[4] http://www.england.nhs.uk/
[5] http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/
[6] https://did.hscic.gov.uk/

Purpose of this document

This paper aims to provide users with an evidence based assessment of the quality of data held within DID by reporting against those of the nine European Statistical System (ESS) quality dimensions and principles[7] appropriate to this output.

In doing so, this meets our obligation to comply with the UK Statistics Authority (UKSA) Code of Practice for Official Statistics[8], particularly Principle 4, Practice 2 which states:

"Ensure that official statistics are produced to a level of quality that meets users' needs, and that users are informed about the quality of statistical outputs, including estimates of the main sources of bias and other errors and other aspects of the European Statistical System definition of quality".

---

[7] The original quality dimensions are: relevance, accuracy and reliability, timeliness and punctuality, accessibility and clarity, and coherence and comparability; these are set out in Eurostat Statistical Law. However more recent quality guidance from Eurostat includes some additional quality principles on: output quality trade-offs, user needs and perceptions, performance cost and respondent burden, and confidentiality, transparency and security.
[8] UKSA Code of Practice for Statistics: http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html

Assessment of statistics against quality dimensions and principles

Relevance

This dimension covers the degree to which the product meets user need in both coverage and content.

The initial need for DID was highlighted by the National Audit Office report, *Managing high value capital equipment in the NHS in England*[9]. This report recommended that data was required on the use of high value equipment used by the NHS to detect and diagnose cancer. It was also identified that such a data set could be designed to allow monitoring of other public health initiatives. As a result the data set was designed to allow the monitoring and reporting of imaging activity across the NHS; GP direct access usage; and to enable linkage to other data sets such as the cancer registration database or Hospital Episode Statistics to provide a wealth of information about care pathways for commissioners, researchers and the NHS community.

DID data is used by a range of users, feeding into national publications and also being used to inform policy and academic research. Some of the main users of DID data are as follows:
- NHS England – DID data is used to produce the monthly DID publication and to inform policy in this area
- Cancer registry – DID data is used to investigate the use of imaging in the diagnosis and treatment of cancer
- Local trusts – DID data is used by local trusts to monitor performance and activity in their own and other areas
- Public and private sector organisations whose use of the data leads to benefits for patients and the public, such as Monitor, systems suppliers, charities etc.
- General Public – DID data is available to the general public through the publication produced by NHS England to provide information on the use of Diagnostic Imaging in the NHS

The data is collected through extracts from local Radiological Information Systems (RIS) and is submitted directly to the HSCIC through a secure website by submitting trusts. Data is collected from all English providers carrying out NHS commissioned imaging and comprises 18 fields, collecting information on the demographic description of the patient; the dates the test was requested, carried out and reported on; the location at which the test was performed; and SNOMED or NICIP codes that describe the test itself. The full list of fields collected can be seen overleaf. Mandatory fields cannot be blank and must contain a valid field. Required fields may be blank, but should contain information if it is available. Fields marked M* denote that at least one of the fields marked in this way must be provided. The 18 fields collected were identified as the minimum data set required to meet the user needs identified above, particularly around reporting on imaging activity and linkage to the Cancer Registry data.

---

[9] http://www.nao.org.uk/wp-content/uploads/2011/03/1011822.pdf

Table 1: List of fields in DID

M/R	Field Name
M*	NHS NUMBER
R	NHS NUMBER STATUS INDICATOR CODE
M*	PERSON BIRTH DATE
M*	ETHNIC CATEGORY
M*	PERSON GENDER CODE CURRENT
M*	POSTCODE OF USUAL ADDRESS
M*	GENERAL MEDICAL PRACTICE CODE (PATIENT REGISTRATION)
M	PATIENT SOURCE SETTING TYPE (DIAGNOSTIC IMAGING)
R	REFERRER CODE
R	REFERRING ORGANISATION CODE
R	DIAGNOSTIC TEST REQUEST DATE
R	DIAGNOSTIC TEST REQUEST RECEIVED DATE
M	DIAGNOSTIC TEST DATE
M	IMAGING CODE (NICIP)
M	IMAGING CODE (SNOMED-CT)
R	SERVICE REPORT ISSUE DATE
M	SITE CODE (OF IMAGING)
M	RADIOLOGICAL ACCESSION NUMBER

DID contains information on tests carried out from April 2012 onwards, and is submitted to the HSCIC on a monthly basis. Providers have three months to submit data for a given month, to allow for updates to the record over time. For example, a test may occur in April, but the report of the test may not be issued until May. The three month rolling period allows trusts to include information such as this in future submissions (up until the end of July in the example given). It is also possible for trusts to request an additional 6 month bypass, whereby they can continue to submit and resubmit data up to 6 months after submission

Feedback on the utility of DID is encouraged, and the HSCIC engages with users through a number of mediums including regular stakeholder meetings; receipt and answer of adhoc enquiries through a dedicated mailbox; and regular interaction with submitters of data to drive up data quality and understand barriers to data submission.

Accuracy and reliability

This dimension covers the statistics proximity between an estimate and the unknown true value.

Accuracy

DID contains record level data submitted by provider, with one record created per examination. The data is submitted through a secure website that carries out a number of validation tests upon submission. Details of these validations can be found in the DID guidance documentation[10]. If a record fails these validation tests the submission may be rejected or the user will be presented with a warning, depending on the severity of the issue. In either case, the submitter is able to act on these warnings and rejections and make multiple submissions, in order to resolve any errors.

Once data has been received by the HSCIC further validation is carried out through a variety of data quality checks. Where there are concerns about data quality we contact providers directly so that any issues with local data extraction processes can be addressed for a future submission. Details of issues which are identified and followed up with Providers by the HSCIC are given below.

Local knowledge, or other comparative data sources, may be required to distinguish changes in volume between reporting periods that reflect changes in service delivery from those that are an artefact of changes in data quality. Such issues should be kept in mind when viewing time series analysis since year-on-year changes may sometimes be a product of shortfalls in earlier years and should not automatically be interpreted as trends in treatment practice or activity.

Submitters have a rolling three month window in which to submit data, and can make multiple submissions during this period in order to ensure that the dataset holds the most accurate and complete record. Upon request, a submitter may be granted an additional bypass to resubmit data up to 6 months after the test date in order to ensure the data is accurate.

When the system encounters an update to an existing record, the existing record is moved to an archive so that only the latest version of the record is present in the live data set.

Where one of the 18 data items collected is classed as required and is left blank, no attempt is made by HSCIC or NHS England to estimate what the missing value should be.

Data quality information for each month is published as part of the NHS England monthly publication, including information on the completeness and consistency of the data in that publication[11].

Reliability/known data quality issues

Missing or Late submissions

Data is collected from all providers performing NHS Commissioned activity recorded in provider systems. Efforts are made to ensure that all eligible providers submit data, but in some cases providers may miss submissions. In addition there may be some providers who should submit data to DID that are not currently doing so. Information on the number of providers for which data was submitted to DID each month can be found in the NHS England publication produced each month.

In addition to the possibility of data not being submitted in a given month, there is variation in the number of updates to data between providers. DID allows data to be submitted in a three month rolling period, to allow records to be updated as more information becomes available, and to allow inaccuracies in the data to be corrected. In some cases, however, submitters do not utilise this three

---

[10] https://did.hscic.gov.uk/Main/Guidance
[11] http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/

month period as intended and instead submit data only once – often at the end of the window. This can mean that records are not updated to include the latest information available, and can increase the impact of a missed submission for a given month.

In addition, it is possible for a submitter to submit data up to 6 months after the date of test, through acquiring a bypass. This is usually to allow the correction of data quality issues; however in some cases submitters have utilised this functionality in order to submit data after the three month deadline, due to developing a backlog of submissions. This affects a small number of providers but will impact on NHS England publications as the data is submitted after the period has been published in NHS England provisional reports. The data will be available in DID once it has been submitted, and will also be included in the annual refresh of data which is published as final data for each financial year period. Where submitters fall behind in this way the HSCIC works closely with them to ensure that normal submissions can resume as soon as possible.

Use of wrong site codes when submitting data

Data is reported by Provider, which is derived from the site code listed in the data submission. In some cases, however, a site may have a number of different site codes to account for services provided by different providers. For example, Leeds General Infirmary has a number of valid provider site codes including RGT1G, RY601, RGD03 and RR801. These codes correspond to four different provider organisations which supply services within Leeds General Infirmary – RGT, RY6, RGD and RR8.

When submitting data it is possible for the submitting organisation to use the wrong site code, causing the activity to be attributed to the wrong provider in NHS England and iView reports. Where submitter-provider mis-matches occur the HSCIC takes steps to contact the submitter affected, to inform them of the issue and to hopefully ensure the issue is corrected in future submissions. In some cases the mismatch may be legitimate, such as when trusts have merged or are in the process of transferring services.

Duplication records

Each record submitted to DID has to have a Radiological Accession Number (RAN). This is generated by the RIS and should be unique to each examination recorded in the RIS. It is not possible for multiple records to have the same RAN in a single submission as the file will be rejected. . It is still possible, however, for duplicate records to be submitted. These fall into two categories:

- Within submitter duplicates – this applies to records which have a distinct RAN within a submitter but all other details about the examination are the same. This may indicate that the record is a duplicate, although it is also possible that the same examination was done to the same patient on the same day (note that only date of test is collected and not time of test). If a high number of within submitter duplicates are identified, either within a submission or across the last three months of data from the same submitter the HSCIC will contact the submitter to determine whether these records have been provided in error, or whether the data is representative of the service provided, and to encourage correction where needed in future submissions.

- Across provider duplicates – this issue applies to records which are in different imaging providers (based on the site code submitted), but where all other elements of the record are the same (The RAN may or may not be the same). This may indicate that the same record has been submitted by more than one organisation in error, although it is also possible for the same exam to be carried out on the same patient at two different providers on the same day. It is likely that this situation will occur where one provider has commissioned another provider to perform the imaging examinations but both are submitting the activity to DID. The DID submitter guidance states that there should be a local agreement on which provider should be responsible for submitting data to DID. This type of duplicate is more complex to identify as it is not limited to a single submission or a single submitter. HSCIC are working to implement at standard monitoring routine for across provider duplicates in 2015. If, however, HSCIC are

alerted to an across provider duplicate issue through use of the data we will contact the submitters responsible to determine the reason for this, and to encourage correction where needed in future submissions.

The impact of duplicate records submitted in error is that data for affected providers may be over counted, rather than presenting a true picture of activity.

Completeness of records
A number of fields submitted to DID are mandatory and therefore cannot be left blank. Some fields, however, are required, meaning that they are only submitted where relevant and can be left blank. Submitters should aim to provide a complete record with data provided for all fields by the end of the three month rolling period. In some cases, however, required fields are not completed by the end of the period, which can reduce the utility of the data. This is often linked to the issue mentioned above, whereby providers submit data only once, and do not update records as more information becomes available.

For example, the report issue date may not be available when the record is first submitted. If the record is not updated in a later submission then the report issue date will remain blank, preventing the data from being used to calculate time waited between test date and report issue and other measures of activity. Poor completeness of required fields should be considered when using and interpreting DID data, and may have an impact on published data, leading to under counting of activity in affected providers.

Where records are consistently missing data from required fields the HSCIC will contact the submitter to understand the reasons for this, and to encourage increased data completeness in future submissions. As part of the monthly publication produced by NHS England[12], data is provided on the completeness of key fields within each reporting period.

Revisions and Archive errors
The DID system is designed to allow the resubmission of records in order for submitters to increase completeness and quality of the data. When a RAN is received in a submission which has previously been submitted to DID, the older record is archived and the new record replaces the old, so that the most recent picture of activity is captured in DID.

In some cases, however, records are archived and replaced with records where at least one key field that would not be expected to change when a record is updated (such as NHS number or test date) has been changed. This kind of update may be to resolve an error in the initial record, but could also indicate that a RAN has been re-used for a new test, rather than a new RAN being generated. It could also indicate that an incorrect identifier is being used as a RAN, such as an event ID, where an event could include a number of examinations which would all have the same event ID.

Instances where this has occurred are known as archive errors. In particular, there is cause for concern when records that are over 6 months old are archived and replaced with a new record with different key fields (This can only occur if the new record has a date of test which has been changed to a date within the last three months). Where a submitter is seen to be generating large numbers of archive errors the HSCIC will contact the submitter to investigate the reasons for this, as it may lead to an undercount of historical activity and result in a change to data that has been published as final data.

Repeated NHS Numbers
A patient may be subject to a number of examinations over time, and each examination will generate a new record in DID. It is, therefore, to be expected that the same NHS number would be recorded multiple times in DID. In some cases, however the same NHS number may be associated with a

---

[12] http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/

11

number of different dates of birth. Although this may be as a result of inaccuracies in the recording of date of birth for a patient, this may also be an indication of the same NHS number being attributed to multiple patients. This is particularly likely in cases where a large number of distinct dates of birth are attributed to the same NHS Number. Evidence suggests that this can happen if a trust has to manually manipulate their extract using software such as MS Excel in order to de-duplicate or format the RIS output to meet the submission requirements of DID.

HSCIC routinely check for occurrence of this and where a submission is identified that contains records with a number of different dates of birth for the same NHS number, the HSCIC will contact the submitter to investigate the reasons for this, and to rectify the issue where relevant in future submissions. Information on occurrences of this issue can also be found in the data quality spreadsheets which accompany the monthly NHS England publication. If not resolved these issues could have an impact on any person related measures, including data linkage to other datasets (such as HES) which rely on NHS Number.

Use of default codes
The DID dataset is validated at source to prevent the submission of invalid values in all appropriate fields. Default codes are permitted as in some cases a valid code may not be applicable or available. While a small level of default codes is to be expected, in some cases organisations may submit a high level of default codes, reducing the utility of the data. One example is the GP practice code of the patient, from which the Clinical Commissioning Group (CCG) is derived. Where default codes are used, it is not possible to attribute the activity to the commissioning CCG, leading to shortfalls in reporting for the affected CCGs.

 Where the HSCIC becomes aware of an organisation providing high numbers of default codes this will be investigated with the submitting organisation to understand the reasons for this, and to implement improvements in future submissions where necessary.

Use of invalid postcodes
Postcodes submitted to DID are validated on submission to ensure they are correctly formatted. There is currently no validation to ensure that postcodes submitted are valid. If a submitter provides invalid post code data this could have an impact on geographic mapping of data. Postcode is used to derive Lower Super Output Area (LSOA) using nationally available postcode reference data. HSCIC plans to monitor the submission of postcodes where no mapping to LSOA can be achieved and contact submitters who include a significant proportion of postcodes which could not be mapped in their submissions, in order to improve the quality of postcodes in DID.

Local system issues

DID contains data from provider organisations' RIS systems. The primary purpose of such systems is to support the provision of patient care and not to provide data for secondary uses, such as national reporting. The configuration of local systems may affect the accuracy of extracts used for DID submissions and may contribute to the issues described above. In many cases providers are at the mercy of system suppliers who may charge a fee to make changes to the RIS configuration.

Timeliness and punctuality

Timeliness refers to the time gap between publication and the reference period. Punctuality refers to the gap between planned and actual publication dates.

Data is submitted over a three month period based on the date of test. For example a test which occurs in November could be submitted in December, January or February[13].

NHS England publish DID data within a month of the end of the three month rolling submission period, so for the above example data for imaging examinations carried out in November would be published towards the end of March. DID linkage data is then published shortly after this point (Details of submission dates for these publications can be found on the NHS England statistics calendar[14] and the HSCIC publication planner[15]).

When an organisation makes use of the six month bypass period they are able to refresh data up to 6 months after the test date (so November data could be submitted up until the end of May). This data will be included in DID, but will not be included in monthly publications as the relevant publication would have been completed and therefore these publications are classed as provisional. This data would be captured in the Annual refresh which is published by NHS England around October of the following financial year.

Submitters are able to access their data through the iView tool within 24 hours of submission, to allow them to assure their submitted data and to monitor performance in a timely manner. Other iView users are able to access data for all submitters once the NHS England publication for that month has been released (this data is rounded and suppressed in accordance with statistical disclosure control rules).

Accessibility and clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

Accessibility

Submitters of DID data are able to access their data through iView within 24 hours of submission, and are also able to download a copy of their submission file for their records. Data is also available using the iView tool following publication of national figures by NHS England for employees of NHS organisations and NHS affiliated organisations, although disclosure control is applied to this data.

The monthly publication of DID data by NHS England contains a PDF report summarising activity for the period, as well as provider level data tables in excel, including a data quality report giving information on quality metrics for that month.

The annual publication of DID data and associated Technical Report by NHS England provide further detail and context to the analysis. For example, these include complarisons with other data sources and alternative breakdowns such as commissioning organisation.

Data is available to the general public through NHS England publications and specific adhoc requests for data can be submitted if more in depth information is required. In addition data extracts

---

[13] It is also possible for a test which occurred in November to be submitted in November, provided the submission date is after the test date provided.
[14] http://www.england.nhs.uk/statistics/12-months-statistics-calendar/
[15] http://www.hscic.gov.uk/pubs/calendar

are available from the Data Access Request Service[16], provided there is a relevant legal basis for the data to be shared.

Clarity

A full specification of the data contained within DID can be found on the Information Standards Board[17] website. This document includes details of the acceptable values that can be submitted to DID and whether fields are mandatory or required.

In addition guidance is available from the DID website[18] which details how data can be submitted to DID, and the validation which is carried out on each field. This can be used to help users of the data understand the processes involved in the submission and validation of data in the dataset.

Users of iView are also able to access detailed guidance which explains the metrics available to produce aggregate data through this tool, and how these can be used. This includes detailed instructions on the functionality available, and also contains details of the statistical disclosure control methods employed in the presentation of this data.

Coherence and comparability

Coherence is the degree to which data which have been derived from different sources or methods but refer to the same topic are similar. Comparability is the degree to which data can be compared over time and domain.

Coherence

There are two aggregate collections of data which contain data related to that which is held within DID. The first is the Annual Imaging and radiodiagnostics collection (KH12), wherein providers submit an aggregate count of how many imaging events occurred, of these how many were accompanied by an intervention, and whether the examinations were carried out by an imaging or other department. The collection is published by NHS England[19], with further in the relevant data dictionary entry[20]. The data is collected and published annually. As DID matures it is intended that it will remove the need for this collection[21], thereby reducing the burden of data collection on the NHS.

Differences may be expected between the two collections, due to the fact that DID is a record level collection and aggregate data is calculated centrally, whereas KH12 is an aggregate collection calculated at a local level. These differences are explored further in comparisons published by NHS England[22]. There are a number of other reasons why DID data may be different to KH12, in particular:
- KH12 only covers specified modalities (and excludes activity for radiotherapy planning) whereas DIDs covers all records held in the RIS at imaging code level, with modalities derived from reference data centrally.
- KH12 counts imaging of multiple body parts separately, whereas DID may combine these into a single imaging examination (although adjustments may be made to improve comparability).
- There are instances where trusts have missed a monthly DID submission or were only able to make a partial submission within the submission window, however this missing activity will be included in the annual aggregate count submitted in KH12.

---

[16] http://www.hscic.gov.uk/dars
[17] http://www.isb.nhs.uk/library/standard/233
[18] https://did.hscic.gov.uk/Main/Guidance
[19] http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostics-waiting-times-and-activity/imaging-and-radiodiagnostics-annual-data/
[20] http://www.datadictionary.nhs.uk/data_dictionary/messages/central_return_forms/hospital_aggregated_statistics/kh12/kh12_fr.asp?shownav=1
[21] https://www.engage.england.nhs.uk/consultation/imaging-radiodiagnostics-statistics/user_uploads/consultation-document.pdf
[22] http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/

A second aggregate collection is carried out on a monthly basis regarding waiting times for diagnostic tests. The Monthly Diagnostic Waiting Times and Activity collection(DM01), collects data on 15 key diagnostic tests and has been published since January 2006. The collection provides information on performance against the national standard that less than 1% of patients should wait 6 weeks or longer for diagnostic tests. The collection is published by NHS England[23], with further guidance on the data dictionary website[24]. The differences between DM01 and DID are included in the comparisons above and many of the same reasons of scope and methodology apply. In addition, the main focus of DM01 is on waiting times and although its activity counts include planned and unscheduled activity, this may be incomplete.

---

[23] http://www.england.nhs.uk/statistics/statistical-work-areas/diagnostics-waiting-times-and-activity/monthly-diagnostics-waiting-times-and-activity/

[24]

http://www.datadictionary.nhs.uk/data_dictionary/messages/central_return_data_sets/data_sets/diagnostics_waiting_times_and_activity_data_set_fr.asp?shownav=1

Comparability

DID data has been collected since April 2012, and data is published monthly based on the most recent information contained within DID for the latest complete month. Due to the functionality which allows data to be resubmitted up to 6 months after the test date, care should be taken when comparing data over time as historical data is subject to change..The data are re-extracted after 6 months and published in an annual refresh of the data. Archive errors, however, may still cause historical data to change in the dataset after this annual refresh has taken place, as there is currently no mechanism to prevent old records being archived, no matter how old they are. HSCIC are planning to implement a system change to rectify this and prevent data older than 6 months from being archived.

Time series analysis of DID data is also vulnerable to changes in service provision; particularly if services are transferred between provider organisations, and these factors should also be considered when comparing DID data over time.

Over time we have employed a number of revisions in order to improve the quality of data within DID. Any data extracts or publications which predate these revisions will differ from data currently stored within DID.

April 2015
Correction of multiple pseudonymised NHS numbers
In early DID submissions submitters were able to submit NHS Numbers with a leading or trailing space. Validation was implemented which prevented this, but NHS Numbers submitted before the validation continued to contain trailing or leading spaces where these had been submitted. As a result the pseudonymised NHS Numbers created for these records treated the space as a character, and so pseudonymised NHS Number generated would differ from one generated for the same NHS number without the leading or trailing blank. In April 2015 this issue was rectified with the leading and trailing blanks being removed, and pseudonymised NHS Numbers being recalculated for the affected records. This affected around 300,000 records in the live tables, and around 750,000 records in the archive table.

Autumn 2014
Referrer Code validation
Initially no validation was carried out upon submission for the referrer code, meaning that invalid codes could be submitted. In autumn 2014 validation was introduced that would allow the submission of invalid codes, but would convert them to a standard value of 99 so that these could be easily identified. This alteration was also applied to retrospective data

Age Calculation
In autumn 2014 the calculation of age was altered so that the age was calculated based on the date of test, rather than the date of submission. This change was applied retrospectively to all data that had been submitted prior to the change, and means that age is now more comparable across the entire data set, regardless of how long a provider takes to submit their data.

Age band Calculation
Age bands were recalculated in autumn 2014 to reflect the changes to the age calculation. In addition, records with an unknown age had initially been grouped into the highest age band, inflating the total count for this age band. From autumn 2014 these records were instead grouped into a separate "Unknown" group. These changes were applied retrospectively to affect all records in DID and resulted in a small number of records moving between age bands where their date of birth falls between the date of test and the date the record was submitted.

Autumn 2013
NHS Number archiving
Where a record is submitted to DID with an accession number that already exists the archiving process treats this as an updated record. It creates a copy of the original record in the archive and then updates each changed data item in the live record with the value contained in the new one.

When DID was launched NHS number was not being included in this update, therefore when NHS number had changed between records the original NHS number was preserved in the updated live record.. In autumn 2013 this issue was rectified so that the old NHS number would be archived and the new NHS number would remain in the live tables. This change was applied retrospectively to all data in DID.

NHS Number Status Indicator archiving
As with the NHS Number, initially when records were submitted with the same accession number as an existing record but a different NHS Number Status Indicator, the old NHS Number status indicator would be preserved. This issue was resolved in autumn 2013, so that the old value would be archived and the new record would be stored in the live tables. This was applied retrospectively to all data in DID.

Ethnic Category Archiving
As with the NHS Number, initially when records were submitted with the same accession number as an existing record but a different ethnic category, the old ethnic category would be preserved. This issue was resolved in autumn 2013, so that the old value would be archived and the new record would be stored in the live tables. This was applied retrospectively to all data in DID.

NHS Number Status Indicator and Ethnic Category in XML Submissions
When DID collections first commenced, ethnic category and NHS Number Status Indicator fields in XML submissions were not being read by the system. This issue was subsequently fixed and the data was cleaned retrospectively in autumn 2013.

Trade-offs between output quality components

This dimension describes the extent to which different aspects of quality are balanced against each other.

The record level DID data is a rich source of information on activity regarding diagnostic imaging, supporting a variety of local uses as well as informing national publications. The data has large potential to increase our understanding of the patient journey and patient care, particularly when linked to other data, such as Hospital Episode Statistics.

Data is provided in a timely manner after submission, whilst also allowing submitting organisations an opportunity to improve the quality and completeness of their data through resubmission. The HSCIC works closely with submitting organisations and stakeholders of DID to continuously improve the quality of the data contained within DID, and there are a number of avenues through which the data can be accessed.
A number of data items that are considered key to the collection, such as Report Issue Date, or items that help identify an individual and therefore improve linkage, are not mandatory. This is a trade-off to allow incomplete or partial records to be submitted in a timely manner. While this allows for more accurate anonymised activity data to be published and analysed it has a negative impact on linkage, care pathway identification and bespoke analysis such as referral to report turn around times.

Assessment of user needs and perceptions

This dimension covers the processes for finding out about users and uses and their views on the statistical products.

The HSCIC liaises closely with stakeholders to establish user needs, including providing support and assistance to submitters of the data. We work with users of the data to understand how DID data can be implemented and to facilitate access to the data where appropriate, within legal restrictions. In addition, in 2013 submitting trusts were surveyed, to generate an understanding of how burdensome submissions to DID were for trusts. We welcome feedback on the utility of DID data and how it can be applied.

Performance, cost and respondent burden

This dimension describes the effectiveness, efficiency and economy of the statistical output.

DID contains information collected through RIS systems reducing the burden on the NHS through the use of routinely collected administrative data. The collection is approved by ROCR, and may allow the retiring of the KH12 collection, further reducing burden on the NHS.

Confidentiality, transparency and security

The procedures and policy used to ensure sound confidentiality, security and transparent practices.

Data is submitted to DID through a secure collection portal, and is then held in secure servers. Data extracts are generally pseudonymised and identifiable data is only shared when there is an approved legal basis to do so. Data available through iView is at an aggregate level only, and the community view is subject to statistical disclosure control. Submitters can access data which is not suppressed relating to their own submission only.

Please see links below to relevant HSCIC policies:
Statistical Governance Policy (see link in 'user documents' on right hand side of page)
http://www.hscic.gov.uk/pubs/calendar
Freedom of Information Process
http://www.hscic.gov.uk/foi
A Guide to Confidentiality in Health and Social Care
http://www.hscic.gov.uk/confguideorg
Privacy and data Protection
http://www.hscic.gov.uk/privacy

This publication may be requested in large print or other formats.

For further information

www.hscic.gov.uk

0845 300 6016

enquiries@hscic.gov.uk