NHS
Business Services Authority

Statistics at NHSBSA

Quality Assurance of Administrative Data (QAAD) Assessment

Prescription Data

Published: July 2020

Document release note

Quality Assurance of Administrative Data (QAAD) Assessment – Prescription Data

Document details name	Version number	Description
Quality Assurance of Administrative Data (QAAD) Assessment – Prescription Data	v001	Document detailing UKSA QAAD assessment of prescription data.

Revision details revision number	Revision date	Revision description	Page number	Previous page number	Action taken	Addenda / new page

About this document

The NHSBSA is responsible for the collection and processing of prescriptions that have been issued in a primary care setting and subsequently dispensed in the community in England. As part of this processing data are collected from prescriptions including information on but not limited to; the drug prescribed and reimbursed, the organisation the prescription was issued by, the individual prescriber that issued the prescription, the contractor that dispensed the prescription, and patient related information such as NHS numbers. These data form the basis for the production of certain Official Statistics released by the NHSBSA such as Prescription Cost Analysis and General Pharmaceutical Services.

This quality assurance of administrative data (QAAD) report focusses on the data produced as part of prescription processing activities, collectively known as ‘prescription data’, and outlines the process data takes from collection through to the output of a statistical release. It identifies the public interest and data quality concerns in this data and the Official Statistics that they appear in, and from this the level of assurance required.

Chapters:

Table of Contents

Section 1: Introduction

Although the NHSBSA is the sole collector, processor and provider of prescription data there are multiple teams involved in the process. Each team has its own clearly defined role in the overarching process and responsibilities relating to this role. NHS Prescription Services, a division of NHSBSA, is responsible for the collection of prescription data and the processing of this data for the payment of dispensing contractors, as well as the assurance of the quality of this data. NHSBSA Insight Data Warehouse Team are responsible for transforming this raw transactional data into more easily accessible formats and collating this information into the NHSBSA Enterprise Data Warehouse, and any quality checks required to ensure that this is done correctly. NHSBSA Official Statistics Team is responsible for producing the official statistics publications that are based upon prescription data and the quality assurance they require. In this document we will demonstrate how each team ensures the quality of prescription data at each step of the process to get from a prescription form to appearing in an official statistics publication.

Section 2: UK Statistics Authority QAAD toolkit

The assessment of our administrative data has been carried out in accordance with the UK Statistics Authority (UKSA) Quality Assurance of Administrative Data (QAAD) Toolkit. QAAD is the Office for Statistics Regulation's (OSR) regulatory standard for the quality assurance of administrative data that is used to produce Official Statistics. The standard recognises the increasing role that administrative data plays in the production of Official Statistics, and clarifies the expectations of producers of Official Statistics to assure themselves of the quality of this data.

We have followed the QAAD toolkit, as described by OSR, and established the level of assurance we do against prescription data used in our Official Statistic releases. Prescription data has been evaluated against the toolkit's risk and profile matrix (Table 3), reflecting the level of risk to data quality and the public interest profile of the statistics produced from this data.

OSR provides example criteria to help determine the level of public interest or value there may be in Official Statistics produced by an organisation (Table 1).

Table 1: Example criteria for each level of public interest or value

Level of public interest or value	Example criteria
Lower	- Always likely to be a politically neutral subject - Interest is limited to a niche user base - Not economically sensitive - Limited media interest
Medium	- Wider range of use for decision making - Moderate economic and/or political sensitivity - Moderate media coverage
Higher	- Legal requirement, for example, for Eurostat - Economically important/market sensitive - Used to inform decisions about resource allocation, for example, by government - Highly politically sensitive, for example, reflected by Select Committee hearings - Substantial media coverage of policies and statistics - Substantial public health concern

From this we have determined the public interest profile for the Official Statistics that prescription data is produced from to be "medium". These publications have a wide range of use for decision and policy making within the Department of Health and Social Care (DHSC) and NHS England & Improvement (NHSE&I). The Official Statistics also gather moderate media coverage upon publication, being the definitive

and only source of information for issues relating to prescribing and dispensing in the community for England.

OSR also provides example criteria to help determine the level of quality concern there may for the administrative data source(s) used to produce Official Statistics (Table 2).

Table 2: Example criteria for each level of data quality concern

Level of data quality concern	**Example criteria**
Lower	- Single data supplier - Simple data collection process - Well-defined classifications - Clear coding frame - Clear instructions for recording - Validation checks built into data collection system - Validation checks built into statistical producer's system - No performance regime or use of targets - International standards for measurement
Medium	Combination of factors from lower and higher levels with safeguards to moderate the outcomes: - More complex data collection - Use of data for payment by results offset by operational checks - Audit established: internal, financial, clinical, sample/statistical - External oversight e.g. by regulators - Multiple data providers offset by use of quality indicators
Higher	- Multiple data supply and/or collection bodies - Complex data collection - Subjective recording of variables - Manual recording and/or coding - Lack of consistency in coding - Lack of clarity in classification systems - No audit of administrative system within operational processes - Over-reliance on system checks - Performance management regime/use of targets - Lack of external oversight

We have determined that prescription data used to produce our Official Statistics has a "medium" level of data quality concern. This is due to the complexities of processing prescriptions and the subsequent data collection processes, alongside the existing audits of the data. We will go into further depth about these processes and audits later in the document.

Having determined the level of quality concern for this data source and the public interest profile of resulting Official Statistics we can use the OSR QAAD risk and profile matrix to determine the level of assurance that we are required to produce (Table 3).

Table 3: UKSA quality assurance of administrative data (QAAD) risk and profile matrix

Level of risk of quality concerns	**Public interest profile**		
	Lower	**Medium**	**Higher**
Low	Statistics of lower quality concern and lower public interest [A1]	Statistics of low quality concern and medium public interest [A1/A2]	Statistics of a low quality concern and higher public interest [A1/A2]
Medium	Statistics of medium quality concern and lower public interest [A1/A2]	Statistics of medium quality concern and medium public interest [A2]	Statistics of medium quality concern and higher public interest [A2/A3]
High	Statistics of higher quality concern and lower public interest [A1/A2/A3]	Statistics of higher quality concern and medium public interest [A3]	Statistics of higher quality concern and higher public interest [A3]

The toolkit outlines four specific areas for assurance that this report will focus on in turn. These are:

- operational context and administrative data collection
- communication with data supply partners
- quality assurance principles, standards and checks applied by data suppliers
- producer's quality assurance investigations and documentation

In the assurance of our data source, we have chosen to give a separate risk and profile matrix score for each of the four areas of assurance. This will allow us to focus our investigatory efforts on areas of particular risk or interest to our users (Table 4).

The QAAD toolkit sets out four levels of quality assurance that may be required of a dataset:

- A0 – no assurance
- A1 – basic assurance
- A2 – enhanced assurance

- A3 – comprehensive assurance

The UK Statistics Authority states that the A0 level is not compliant with the Code of Practice for Statistics. The assessment of the assurance level is in turn based on a combination of assessments of data quality risk and public interest. The toolkit sets out the level of assurances required as follow:

Level A1 – basic assurance
The statistical producer has reviewed and published a summary of the administrative data quality assurance (QA) arrangements.

Level A2 – enhanced assurance
The statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance.

Level A3 – comprehensive assurance
The statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit and published detailed documentation about the assurance and audit.

Within the UK Statistics Authority QAAD – setting the standard documentation it states:

"Quality assurance of administrative data is more than simple checking that the figures add up. It is an on-going iterative process to assess the data's fitness to serve their purpose. It covers the entire statistical production process and involves monitoring data quality over time and reporting on variations in that quality. Post collection quality assurance methods, such as data validation, are an important part of the quality assurance process, but can be of limited value if the underlying data are of poor quality. The Authority encourages the application of critical judgement of the underlying data from administrative systems before the data are extracted for supply to the statistical production process. … producers need to: investigate the administrative data to identify errors, uncertainty and potential bias in the data; make efforts to understand why these errors occur and to manage or, if possible, eliminate them; and communicate to users how these could affect the statistics and their use."

The outcome of our assessment then determines the level of assurance and documentation required to inform people about the data QA arrangements in place for the administrative systems from which our statistics are sourced.

The results of those assessments suggest an assurance level of A2 (enhanced assurance) is required.

Section 3: Assessment and justification against the QAAD risk and profile matrix

Table 4: QAAD risk and profile matrix assessment of Prescription administrative data.

	Low (A1)	Medium (A2)	High (A3)
Operational context and administrative data collection		A2: Enhanced assurance.	
Communication with data supply partners		A2: Enhanced assurance.	
Quality assurance principles, standards and checks by data supplier		A2: Enhanced assurance.	
Producers quality assurance investigations and documentation		A2: Enhanced assurance.	

We have determined that the level of quality concern for prescription data, and public interest profile relating to Official Statistics produced from this data is at a medium level. Across all four areas of assurance we focus on in this report an A2 – enhanced assurance level has been completed.

In the following section we will address each of these four areas in turn and describe how we meet this level of assurance.

Section 4: Areas of quality assurance of administrative data (QAAD)

4.1 Operational context and administrative data collection (QAAD Toolkit matrix score A2)

This relates to the need for statistical producers to gain an understanding of the environment and processes in which the administrative data is being compiled, and the factors that might increase the risks to the quality of the administrative data.

NHS Prescription Services, a division within the NHSBSA, is responsible for collecting, processing, and subsequently reimbursing dispensing contractors for prescription items dispensed, and other essential and advanced services provided to NHS patients.

Data collection

Prescription data are captured from prescription forms, both paper and electronic. Electronic prescribing via the Electronic Prescription Service (EPS) now accounts for over 75% of all prescription items, and is therefore the primary mechanism that NHS Prescription Services uses to collect prescription data. Dispensing contractors can submit information to NHS Prescription Services via EPS from any point after the prescription items have been dispensed to the patient, but this should be submitted within the same month as they were dispensed.

For paper prescriptions, contractors submit a collated batch at the end of the month that the prescriptions relate to NHS Prescription Services via registered courier for processing. These paper forms are then scanned to collect digital images and passed through intelligent character recognition (ICR) to extract data to pass to the transactional systems.

Data from EPS messages and captured from paper prescription forms are passed to the main transactional system, Capacity Improvement Programme (CIP) which applies the pricing rules to each prescription item. As part of this some items are presented to an operator if an exception rule is met, such as an unusual quantity or if a product is not recognised and cannot be processed. This introduces a manual aspect to this process and increases the risk of error occurring. The prescriptions most likely to be presented to an operator and require manual intervention are handwritten paper forms. Anecdotally these forms are mostly issued by nurse prescribers, hospital trusts and dental prescribers, therefore errors will more likely be biased towards these types of prescribers.

Once processing has been completed by NHS Prescription Services, extracts from the transactional system (CIP) as well as all other source systems that are used for identifying other aspects relating to prescription data such as drug information and prescriber information are taken. These extracts are passed to the Data Warehouse

Team, a part of the Insight directorate within the NHSBSA, who then extract, transform and load (ETL) the data into the NHSBSA Enterprise Data Warehouse. As part of this transform and load process business logic is applied to the raw transactional data to make it more usable and understandable and additional verification checks carried out to ensure that all data has been collected, if any checks fail, these are investigated immediately and errors rectified before the data is made available to statistics producers and users of NHSBSA Information Systems.

A full prescription data dataflow and process map is provided in annex A.

Strengths

- Dispensing contractors are required to submit their prescriptions to NHS Prescription Services in order to be reimbursed. Therefore coverage can reasonably be expected to be close to 100%.
- Data are all collected and processed in the same manner from a single source, NHS Prescription Services.
- Prescription data submitted by EPS has a data capture accuracy of 99.99% and over.
- The majority of data are collected via EPS with the proportions of this expected to increase in the future.

Weaknesses

- Relative timescales from receipt of prescriptions to release of data. Data are available approximately six weeks after the end of the month to which the data relate.
- Prescription data are captured as a by-product of processing prescriptions for payment. Therefore some things are not captured reliably as this is not needed for reimbursement, such as the actual medicinal product that has been dispensed to a patient.

4.2 Communication with data supply partners (QAAD matrix Toolkit score A2)

This relates to the need to maintain effective relationships with suppliers (through written agreements such as service level agreements or memoranda of understanding), which include change management processes and the consideration of statistical needs when changes are being made to relevant administrative systems.

Communication with NHS Prescription Services

As all data is collected and processed by a division of the same organisation there are established robust lines of communication and working relationships between statistics producers and the relevant people within NHS Prescription services, such as the Quality and Accuracy Improvement Team, Medicines Data Team, and Organisational Data Team. Any issues with data that have been identified at source can be quickly communicated to the relevant team via email, telephone, face-to-face

meeting or less formal methods such as instant messenger for example Microsoft Teams.

NHSBSA statisticians are involved in the change control process that relates to information systems to ensure that as far as possible statistical needs are taken into account in any decision-making around changes in the data collected or its format.

Communication with NHSBSA Data Warehouse Team

The Official Statistics Team sits in the same directorate within the NHSBSA as the Data Warehouse Team. Statistical staff have strong links with Data Warehouse staff, regularly using formal and informal communication routes to raise any issues identified with statistical data. JIRA, a project management and software delivery tool, is used to track progress with these issues, and also to log requests for creation of any 'new' data within the NHSBSA enterprise data warehouse along with full requirements for this.

Data is transferred between the Data Warehouse Team and the Official Statistics Team via database tables stored within a secure cloud based environment, only accessible via NHSBSA devices with valid credentials. Table access within the NHSBSA enterprise data warehouse is managed by the Data Warehouse Team, with access only granted to specific tables for each user, based upon the requirements to fulfil their job role.

NHSBSA statisticians are also consulted about the data content of any new information reporting systems when they are produced, and are involved in the change control process relating to existing reporting systems.

Strengths

- Multi-channel communication between the different teams within the NHSBSA that are responsible for the creation, storage and reporting of prescription data.
- Fully defined requirements for data that are held within the NHSBSA enterprise data warehouse, and those subsequently used in statistical publications.
- All data collected, processed and stored by the same organisation, with different teams regularly collaborating and having an understanding of each other's operating procedures.

4.3 Quality assurance principles, standards and checks by data supplier (QAAD Toolkit matrix score A2)

This relates to the validation checks and procedures undertaken by the data supplier, any process of audit of the operational system and any steps taken to determine the accuracy of the administrative data.

Quality assurance carried out by NHS Prescription Services
NHS Prescription Services carries out a series of automated and manual quality assurance processes to verify the accuracy of data captured. Any prescription items identified by the automated QA processes are presented to an operator to manually verify and input any data that has been flagged as erroneous.

Each month a statistically valid random sample of 50,000 prescription items are reprocessed to identify any errors that were made in the reimbursement of the item or in the information captured for reporting purposes. This process is internally audited and results are expected to exceed 99.50% at all times. Results are published on the NHSBSA website for transparency, and for the twelve month period from January 2019 to December 2019 had an average value of 99.75%.

There is also a dedicated Quality and Accuracy Team within NHS Prescription Services who provide feedback to operators regarding errors made, providing guidance around best practice. Operators receive a monthly feedback check of 100 prescription items that is separate from the 50k check used to determine overall divisional accuracy. Any errors made are reported back to the operator for information purposes.

Quality assurance carried out by NHSBSA Data Warehouse Team
The NHSBSA Data Warehouse Team have a comprehensive set of verification checks that are carried out during the monthly extract, transform and load (ETL) process to populate the NHSBSA enterprise data warehouse administrative data source. These checks ensure that the data that is being extracted from the NHSBSA transactional systems are valid and complete. Any errors identified by these checks are investigated and resolved immediately before the load process is allowed to continue.

The code that applies business logic to the transactional data as part of the ETL process has been developed and quality assured by Data Warehouse staff and is a mature codebase. All code is version controlled and development work tested and reviewed before being deployed into the production system.

Strengths
- Mature QA processes in place in NHS Prescription services, with results available publicly for scrutiny.
- Culture of accuracy in place throughout the NHSBSA.
- Checks carried out at an operational level and at a reporting level.

Weaknesses
- Certain data, such as organisational data is not verified against other sources, such as NHS Digital.
- Even with QA processes due to manual nature of some work there will always be a number of errors.

- Main concern for NHS Prescription Services is correct payment of dispensing contractors, rather than capture of reporting information. However, this may lend itself to ensuring the quality of captured information.

4.4 Producers quality assurance investigations and documentation (QAAD Toolkit matrix score A2)

This relates to the quality assurance conducted by the statistical producer, including corroboration against other data sources.

Due to the nature of the NHSBSA enterprise data warehouse and its basis as the main source of data for all NHSBSA information systems, some datasets that are provided to the Official Statistics Team are produced by code written and maintained by the Data Warehouse Team. This is done in situations where the size of the datasets and the process required to produce them is deemed to have an impact on the production environment and affect other users of the systems. In these circumstances the code used to produce these data sets is written by a Data Warehouse and Repository Developer and quality assured by a Data Warehouse and Repository Developer or higher within the team.

Automated QA processes

Each statistical publication that is on-boarded or created by the Official Statistics Team is developed into a reproducible analytical pipeline (RAP). As part of this, code is written to help automate, and expedite the quality assurance process of each publication. Validation checks are written to assess the accuracy and validity of the 'raw' data that the publication is based upon; further checks are also put in place to ensure that the outputs from the RAP are as expected and within certain tolerances from previous publications. The results from these checks are then written to CSV files to be passed to a Senior Statistical Officer (SStO) to carry out independent quality assurance. The code written to create these automated QA checks is also peer reviewed, with unit testing to be developed as the RAP matures. The code is also published to the NHSBSA GitLab to allow it to be viewed and interrogated by the public.

Manual QA processes

Once the publication has been produced by the RAP, all files including; a statistical summary, aggregated tables, and any additional tables, are handed to another SStO, equivalent, or higher within the NHSBSA to carry out further manual checks. These include checking formatting and other typographical issues. Checks are also carried out against other published sources of information such as the English Prescribing Dataset, a monthly administrative data feed provided by NHSBSA Information Services. Publications are sent for final sign off by the Lead Official for Statistics at the NHSBSA before being made public.

The NHSBSA is also collaborating with other government departments and areas of the Government Statistical Service (GSS) to review our processes and help develop our statistical capability.

Strengths

- No statistical output is released without having been reviewed and signed off by at least two senior members of the Official Statistics Team.
- Portions of QA automated to allow checking of large volumes of data.

Weaknesses

- Some errors might only be noticeable by someone with extensive knowledge of prescriptions data.
- Maturation of automated QA within RAP needed to find approach that works best for the Official Statistics Team.

Section 5: Summary

The NHSBSA considers the main strengths of the prescriptions dataset to be:

- Dispensing contractors are required to submit their prescriptions to NHS Prescription Services in order to be reimbursed. Therefore coverage can reasonably expected to be close to 100%.
- Data are all collected and processed in the same manner from a single source, NHS Prescription Services.
- Prescription data submitted by EPS has a data capture accuracy of 99.99% and over.
- The majority of data are collected via EPS with the proportions of this expected to increase in the future.
- All data collected, processed and stored by the same organisation, with different teams regularly collaborating and having an understanding of each other's operating procedures.
- Mature quality assurance processes in place in NHS Prescription services, with results available publicly for scrutiny.
- Culture of accuracy in place throughout the NHSBSA.
- Checks are carried out at an operational level and at a reporting level.
- No statistical output is released without having been reviewed and signed off by at least two senior members of the Official Statistics Team, and finally by our Lead Official for Statistics.

We believe the main current limitations of this data source are:

- Relative timescales from receipt of prescriptions to release of data. Data are available approximately six weeks after the end of the month to which the data relate.
- Prescription data are captured as a by-product of processing prescriptions for payment; therefore some things are not captured reliably as this is not needed for reimbursement, such as the actual medicinal product that has been dispensed to a patient.
- Certain data, such as organisational data is not verified against other sources, such as NHS Digital.
- Even with QA processes due to manual nature of some work there will always be a number of errors. In the event these are significant errors we will follow our Revisions and Corrections policy for statistics.
- Some errors might only be noticeable by someone with extensive knowledge of prescriptions data.

In constantly seeking to improve our data sources, we will be taking next steps to investigate these limitations, and these will be communicated to users in future quality assurance of administrative data (QAAD) reports. However, currently Official

Statistics produced by the NHSBSA from this data source are assured to level A2 (enhanced assurance) as per the UKSA QAAD toolkit, as we have evaluated the administrative data QA arrangements and published a fuller description of the assurance

Contact us

Feedback is important to us; we welcome any questions and comments relating to this document.

Please quote:

'Quality Assurance of Administrative Data (QAAD) Assessment – Prescription Data'

You can contact us by:

Email: nhsbsa.statistics@nhs.net

Telephone: 0191 203 5050

You can also write to us at:

NHSBSA - Statistics
NHS Business Services Authority
Stella House
Goldcrest Way
Newburn Riverside
Newcastle upon Tyne
NE15 8NY

END