

29 January 2015

## Quality Assurance of Administrative Data

The UK Statistics Authority has today published a Regulatory Standard for the Quality Assurance of Administrative Data.<sup>1</sup>

The Authority produced this Standard in response to concerns about the quality of administrative data that emerged during its assessments of statistics on police recorded crime.<sup>2</sup> The Standard recognises the increasing role that administrative data are playing in the production of official statistics and clarifies the Authority's expectations for what producers of official statistics should do to assure themselves of the quality of these data.

The Standard encourages risk-based judgement and supports a proportionate approach, recognising that not all administrative data sources are high risk. As a result, the Standard is pragmatic, and the toolkit that supports it provides helpful guidance to statistical producers about the practices they can adopt to assure the quality of the data they receive.

Ed Humpherson, the Statistics Authority's Director General for Regulation, today said:

"The Standard encourages statisticians not to take data at face value, but to be willing to question the data they receive and to convey this scepticism to the users of their statistics. We are grateful for the advice and input we have received from Government statisticians and others in the development of this standard and are confident that statisticians will be capable of implementing the Standard's principles; indeed, many already are."

---

<sup>1</sup> <http://www.statisticsauthority.gov.uk/assessment/monitoring/administrative-data-and-official-statistics/index.html>

<sup>2</sup> <http://www.statisticsauthority.gov.uk/reports---correspondence/correspondence/letter-from-sir-andrew-dilnot-to-bernard-jenkin-mp-150114.pdf>



# **Quality Assurance of Administrative Data**

## **Setting the Standard**

**Version 1 January 2015**

On 31 July 2014 the UK Statistics Authority published an Exposure Draft report [Quality Assurance and Audit Arrangements for Administrative Data](#).

This report presents the Authority's regulatory standard that it will use to assess statistics derived from administrative sources against the *Code of Practice for Official Statistics*. It is published alongside the [Administrative Data Quality Assurance Toolkit](#).

The Authority will release further supporting information to guide both statistical producers and non-statistical officials in using administrative data for official statistics:

- Emerging case studies of good practice
- Guidance for policymakers when using statistics
- Response to consultation

## Introduction

1. This regulatory standard confirms the quality assurance arrangements that are required for statistics that are compiled using administrative data to comply with the *Code of Practice for Official Statistics*<sup>1</sup>. The *Administrative Data Quality Assurance Toolkit* (QA Toolkit) is the mechanism that the Authority will use to determine compliance. The Toolkit is built around the Quality Assurance Matrix (QA Matrix) which presents the levels of assurance for four areas of practice related to the quality assurance of official statistics and of the administrative data used to produce them.
2. Our assessors will:
  - Provisionally determine their view of the appropriate level of assurance required for a set of statistics.
  - Request from statistical producers:
    - a) an explanation of their judgments of the required level of assurance for a set of statistics;
    - b) evidence to support the rationale for these decisions;
    - c) evidence of the actions, and the rationale for deciding upon those actions, that the producer has taken to achieve compliance with the chosen level of assurance; and
    - d) evidence which demonstrates that the statistical producer has embedded practices for keeping its quality assurance arrangements under review.
  - Identify whether the submitted evidence complies with the *Code of Practice*.
  - For instances of non-compliance, require statistical producers to take action to meet the standard of the *Code of Practice*.
3. This paper sets out the basis of the QA Toolkit in relation to the *Code of Practice*. It also addresses a number of the issues raised by respondents to the Exposure Draft, in relation to the definition and nature of administrative data, the need for repeated and

<sup>1</sup> <http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html>

ongoing evaluation of data quality, the complex nature of data supply relationships, and the role of systematic external audit.

4. The QA Toolkit recognises that a proportionate approach is appropriate, whereby some statistics will need greater levels of assurance than others. The levels can be determined by an evaluation of: (i) the likelihood of quality concerns arising in the administrative data that may affect the quality of the statistics; and (ii) the nature of the public interest served by the statistics.
5. The *Code of Practice* requires statistical producers to provide explanatory information to help users understand the strengths and limitations of official statistics in relation to their use. Producers need to ensure that they have a good appreciation of the impact of the data collection processes on the statistics and that they communicate these issues to those who use the statistics. It is important to critique the data at each stage of the statistical production process.
6. The Authority encourages statistical producers to use this Toolkit routinely as part of their analysis and monitoring of administrative data systems, and to share their findings with users.

## Background

7. On 31 July 2014 the Authority published an Exposure Draft report *Quality Assurance and Audit Arrangements for Administrative Data*. The Exposure Draft presented the Authority's proposed regulatory standard for the quality assurance of administrative data that are used to produce official statistics. We considered the risks associated with the use of administrative data for statistical purposes, identified some examples of existing practices across government in addressing those risks, and presented our proposed mechanisms for statisticians to use when seeking to implement best practice.

## Consultation

8. The Authority invited comments on the proposed approach. In response we received 14 written comments from private individuals, government departments and from statisticians responding in personal and professional capacities. Between September and December 2014 we spoke directly to over one hundred statisticians in government departments across the UK, as well as to some policymakers, public and private sector data suppliers and other regulatory bodies.
9. Overall, the feedback endorsed the proposals and reassured us that the approach was sound. It was clear that there is a common understanding that it is important for statistical producers to gain, and share with users, a fuller understanding of the administrative data that they use to produce official statistics, of the circumstances in which they are produced, and how they are tested and verified. Statisticians told us that they face many challenges when using administrative data, including having limited resources and obtaining data from old, poorly documented data systems. Statisticians also told us of some frustrations in having to react to policy changes rather than being consulted as part of the process. There was also a request for examples of good practice

to be shared between statistical producers. We also heard from dynamic and enthused statisticians determined to implement improved processes and who shared with us good examples of their work.

10. We will publish supporting documents on our website, these will include:
- revised *Guidance for policymakers using statistics*;
  - a series of case studies that demonstrate good practice in aspects of quality; assurance
  - key responses to the consultation and the Authority's response.

## The Code of Practice for Official Statistics

11. The *Code of Practice* encompasses the production, management and dissemination of official statistics – this covers the entire statistical process from the identification of needs, to the decision to collect or compile data, through to providing advice to the user. It requires statistical producers to ensure that administrative sources are fully exploited for statistical purposes, with appropriate safeguards in place<sup>2</sup> and that they:

- Use data based on definitions and concepts which approximate well with those required for the statistics<sup>3</sup>.
- Inform users about the quality of their statistical outputs, including estimates of the main sources of bias and other errors in the data<sup>4</sup> and the quality and reliability of statistics in relation to the range of potential uses<sup>5</sup>.
- Publish supporting metadata which includes information about the quality assurance procedures<sup>6</sup>, quality guidelines<sup>7</sup>, the administrative systems used in the production of official statistics<sup>8</sup> and those with the potential to be so used<sup>9</sup>, and the arrangements for auditing the quality of the data<sup>10</sup>.
- Identify arrangements for allowing statistical staff access to administrative data for statistical purposes and ensure that full account is taken of the implication for official statistics when changes to administrative systems are considered<sup>11</sup>.
- Promote statistical purposes actively in the design of administrative systems in order to enhance the statistical potential of administrative records<sup>12</sup>.
- Maximise opportunities for the use of administrative data, cross-analysis of sources and for the exchange and re-use of data, to avoid duplicating requests for information<sup>13</sup>.

---

<sup>2</sup> Protocol 3 of the *Code of Practice for Official Statistics*

<sup>3</sup> Protocol 3, Practice 2 of the *Code of Practice for Official Statistics*

<sup>4</sup> Principle 4, Practice 2 and Protocol 1, Practice 4 of the *Code of Practice for Official Statistics*

<sup>5</sup> Principle 8, Practice 1 of the *Code of Practice for Official Statistics*

<sup>6</sup> Principle 4, Practice 3 of the *Code of Practice for Official Statistics*

<sup>7</sup> Principle 4, Practice 4 of the *Code of Practice for Official Statistics*

<sup>8</sup> Protocol 3, Practice 5(a) of the *Code of Practice for Official Statistics*

<sup>9</sup> Protocol 3, Practice 5(c) of the *Code of Practice for Official Statistics*

<sup>10</sup> Protocol 3, Practice 5(e) of the *Code of Practice for Official Statistics*

<sup>11</sup> Protocol 3, Practice 5(b) and 5(d) of the *Code of Practice for Official Statistics*

<sup>12</sup> Principle 6, Practice 3 of the *Code of Practice for Official Statistics*

<sup>13</sup> Protocol 3, Practice 3 of the *Code of Practice for Official Statistics*

## Administrative data

12. Administrative data refers to information collected primarily for administrative reasons (not research). This type of data is collected by government departments and other organisations for registration, transactions and record-keeping, usually when delivering a service<sup>14</sup>. Administrative data are often used for operational purposes and their statistical use is secondary.
13. The Exposure Draft outlined some of the benefits and the challenges of using administrative data to produce official statistics. The purpose of this paper is to introduce the regulatory standard rather than set out a definitive description of those merits and risks.
14. The Authority recognises that there are limitations with administrative data and that these can create complications when compiling official statistics. However their use is central to the production of official statistics and the existence of these challenges places a premium on proactive quality assurance to investigate the data, manage identified issues, and clearly communicate any limitations to users.

## Quality Assurance of administrative data

15. Quality assurance of administrative data is more than simply checking that the figures add up. It is an ongoing, 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 judgment of the underlying data from administrative systems *before* the data are extracted for supply into the statistical production process. As with survey data, 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.

## Using the Quality Assurance Matrix

### Practice Areas

16. In an Assessment of official statistics based on administrative data, the Authority will consider the quality assurance of the data across the four practice areas outlined in the QA Toolkit. These practice areas demonstrate the need for the quality assurance of statistics obtained from administrative statistics to extend beyond the checks made by statistical producers on the data they receive. The producers should demonstrate knowledge of the operational context in which the data are recorded, and an understanding of the impact that the motivations of data inputters can have on the data.

---

<sup>14</sup> Definition from the Administrative Data Research Network <http://www.adrn.ac.uk/faq/about-the-data>

They should also have good communication links with data supply partners and understand their partners' data quality processes and standards.

### Four practice areas associated with data quality

Operational context & admin data collection	Communication with data supply partners	QA principles, standards and checks by data suppliers	Producers' QA investigations & documentation
<ul style="list-style-type: none"> <li>• environment and processes for compiling the administrative data</li> <li>• factors which affect data quality and cause bias</li> <li>• safeguards which minimise the risks</li> <li>• role of performance measurements and targets; potential for distortive effects</li> </ul>	<ul style="list-style-type: none"> <li>• collaborative relationships with data collectors, suppliers, IT specialists, policy and operational officials</li> <li>• formal agreements detailing arrangements</li> <li>• regular engagement with collectors, suppliers and users</li> </ul>	<ul style="list-style-type: none"> <li>• data assurance arrangements in data collection and supply</li> <li>• quality information about the data from suppliers</li> <li>• role of operational inspection and internal/external audit in data assurance process</li> </ul>	<ul style="list-style-type: none"> <li>• QA checks carried out by statistics producer</li> <li>• quality indicators for input data and output statistics</li> <li>• strengths and limitations of the data in relation to use</li> <li>• explanation for users about the data quality and impact on the statistics</li> </ul>

### The Quality Assurance Matrix

- The Authority recognises the resource challenges faced by statistical producers and supports a proportionate and pragmatic approach to the assurance of administrative data in relation to each area of practice. The QA Matrix encourages producers to consider the level of data quality issues that could affect their data together with the level of public interest in the statistics.
- The QA Matrix is **not** intended to be used as a check-list or a tick box exercise. Work has been done to develop such lists<sup>15</sup> and they can be useful in developing quality assurance processes. However, over-reliance on a checklist may lead to cases where issues are over-simplified and limit the ongoing development of quality assurance processes. The QA Matrix presents examples of the types of evidence that assessors would expect to see as part of an Assessment, to provide statistical producers with an understanding of the standard required for each level of assurance.
- The assurance levels set out in the QA Matrix are not intended to be interpreted as a RAG status<sup>16</sup> as commonly used in project management. For example level A3 does not necessarily mean that remedial action is needed; instead it could mean that the statistics have ongoing levels of higher public interest and higher quality concerns which would mean that higher levels of assurance would be continual. Remaining at A3 level of

<sup>15</sup>For example see <http://essnet.admindata.eu/WorkPackage?objectId=4257>

<sup>16</sup> RAG (Red, Amber, Green) status is commonly used in project management to visually highlight the risk level of a project, with red showing the greatest risk and green the least.

assurance over time does not represent a failing by the statisticians; it simply demonstrates that a higher level of assurance is necessary for a certain set of statistics.

20. The QA Matrix is not exhaustive. It does not present all the possible types of work that statistical producers could carry out to assure the quality of their data. In many cases they will be undertaking other types of quality assurance work which they should share with assessors. Similarly, statistical producers might recognise that not all elements suggested at a certain level of assurance will be appropriate for their statistics, and have appropriate reasons for making such a judgment. Ultimately it is for each statistical producer to decide how it meets the standard and present this evidence to the Authority.
21. A key element of practice emphasised in the QA Toolkit is that if flaws are found in administrative data, statistical producers should: evaluate the likely impact on the statistics; establish whether the issue can be resolved, or whether there is any other action they can take to mitigate the risks; and determine whether the level of impact is such that users should be notified. It is recognised that often issues discovered through quality assurance are complex and will require time and staffing and financial resources to address. Statistical producers are required to maintain ongoing compliance with the *Code of Practice*. If, in the course of these investigations, a statistical producer discovers a systemic issue in the administrative data that has a substantial adverse impact on the statistics, we encourage the statistical Head of Profession to contact the Authority to discuss appropriate action.
22. The Authority recognises that there are certain circumstances in which regular, systematic external evaluation, audit or inspection of the underlying data is essential to increase both the quality of, and public confidence in, statistics produced from administrative data. For statistics requiring higher levels of assurance, these external evaluations should be regular and repeated. In the absence of regular repeated external scrutiny (for those sets of statistics for which it is appropriate) the statisticians should highlight the deficiency for users and investigate whether there are other sources that could be used.
23. Administrative data underpinning official statistics can be subject to, or feature in, various kinds of audit, depending on their operational context, for example: financial, clinical, and statistical audit in which a sample of existing cases is investigated. These investigations can be used to provide context, and in some cases corroboration of the data quality. It is good practice to investigate whether these other types of audits could provide information to support the statisticians' quality judgments, and to use them if appropriate. It is also recognised that in some cases these types of corroborating information are not available.
24. The steps to be taken by statistical producers need to go beyond a narrow interpretation of 'quality assurance'; they should also encompass the working arrangements and relationships with other agents, particularly data supply partners. These working arrangements can range from straightforward data transfers from a single data supplier, through to more complex large-scale systems. In addition some official statistics are based on administrative data provided by another statistical producer body.



For statistics based on administrative data obtained from outside the Government Statistical Service some examples include:

- **Directly from the organisation that records the data** – in this case the statistical producer should engage with this organisation directly and take into consideration the detail and nature of information received from the data supplier when deciding on appropriate quality assurance. The responsibility for producing information about the quality of the statistics lies with the producer. In the case of large numbers of direct suppliers, the statistical producer might explore other ways of engaging, such as holding meetings to discuss common quality issues and liaising with information governance groups.
- **Via intermediary organisations** – for example local authorities collect data and pass these to the producer – in this case the statistical producer should: 1) have an understanding of the entire data cycle and the quality assurance processes carried out by the original data supplier; and 2) engage with intermediary organisations to understand their quality assurance processes and standards.

For statistics based on administrative data collected by another government department or official statistics producer body:

- In this case both the statistical producer and the supplier department have a responsibility to understand and communicate about the quality of the data. The producer with responsibility for publishing the statistics has the ultimate responsibility for assuring itself about the quality of the underlying data and communicating this to users. However the statistical Head of Profession for the supplier partner body has a responsibility to ensure that it shares its quality assurance information with the statistical producers, including clear information about any limitations or bias in the data, and takes action to address any queries or concerns raised by the statistical producer. In cases where producer bodies share data for statistical purposes, or both work with the same data supply partners, we encourage such producers to work together to develop a better understanding of the quality of the administrative data, sharing intelligence and insight about the data with each other and with users.

## Guide for policymakers

25. As part of the suite of reports published alongside this paper, the Authority will publish *Guidance for policymakers using statistics*. This will build upon Annex A of the Exposure Draft report. As part of the feedback we received to the Exposure Draft, statisticians told us of the need to share with policymakers the benefits and limitations of official statistics when policy decisions are made. Statistical input is an important part of policy making and this guidance will point to critical questions that policymakers can ask about the statistics they use and of the statisticians who produce them.

# **Administrative Data Quality Assurance Toolkit**

**Version 1 January 2015**

## Administrative Data Quality Assurance Toolkit

This toolkit is intended to help statistical assessors review the areas of practice for the quality assurance arrangements of administrative data used to produce official statistics, during the Assessment of official statistics against the *Code of Practice*<sup>1</sup> for possible designation as National Statistics.

The Authority is sharing this toolkit to help explain the nature of the regulatory standard that we are applying when we determine the suitability of practices used by statistics producers for the quality assurance of administrative data.

This guidance can support producers in making their critical judgments that administrative data continue to be suitable to be used in producing statistics, and in providing appropriate documentation for users.

The toolkit comprises:

- a) *Quality Management Actions* (page 3) – a model describing an approach that can be used by statistics producers to satisfy themselves and their users of the ongoing suitability of the administrative data
- b) *QA Matrix* (page 4) – which describes the statistical practices in four main areas for ensuring the suitability of the administrative data quality, with three different levels of assurance as some statistics require a more thorough and detailed approach than others
- c) *Risk/Profile Matrix* (page 9) – which can be used to inform decisions about the level of assurance that is appropriate given the characteristics of the statistics
- d) *Top tips and Helpful Guidance* (pages 10 and 11) – additional guidance material to support judgments about the quality of administrative data

Our assessors will:

- Provisionally determine their view of the appropriate level of assurance required for a set of statistics.
- Request from statistical producers:
  - a) an explanation of their judgments of the required level of assurance for a set of statistics;
  - b) evidence to support the rationale for these decisions;
  - c) evidence of the actions, and the rationale for deciding upon those actions, that the producer has taken to achieve compliance with the chosen level of assurance; and
  - d) evidence which demonstrates that the statistical producer has embedded practices for keeping its quality assurance arrangements under review.
- Identify whether the submitted evidence complies with the *Code of Practice*.
- For instances of non-compliance, require statistical producers to take action to meet the standard of the *Code of Practice*.

---

<sup>1</sup> <http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html>

## Quality Management Actions

Assessors will look for evidence that producers have considered the following quality management actions: *Investigate*: *Manage*: *Communicate*. They represent three types of actions for assuring the suitability of the administrative data and in documenting the findings:

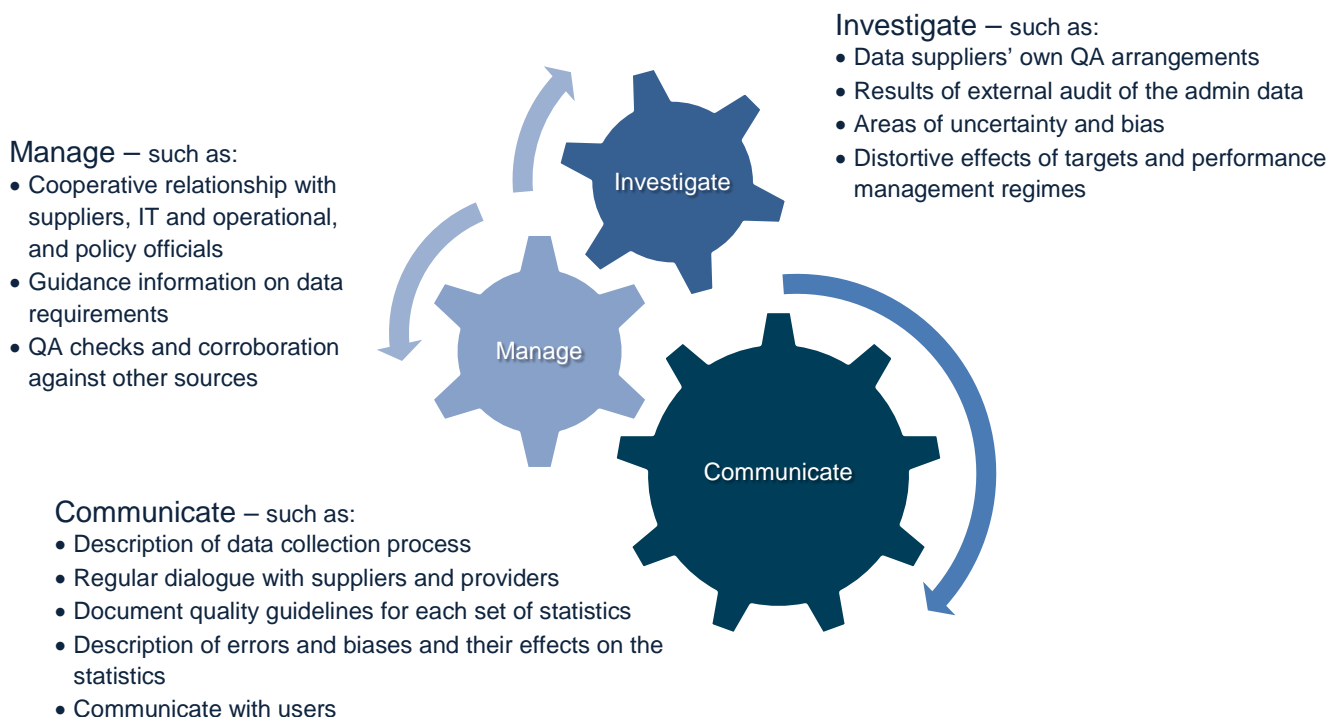
- **Investigate:** Statistics producers should investigate, for example, the types of checks carried out by data collectors and suppliers, as well as the operational circumstances in which the data are produced. They should identify any coverage issues and potential sources of bias in the data collection and supply process.
- **Manage:** Producers should also manage their relationships with suppliers by establishing clear processes for data provision and for managing change. They should also maintain regular quality assurance checks of the data and use other data sources where possible to corroborate their findings.
- **Communicate:** Producers should communicate effectively with their data suppliers and others to ensure users are provided with clear explanations of the strengths and limitations of the data. Producers should work closely with other statistical producers using the administrative data to ensure a common understanding of any quality issues.

As illustrated in the diagram (below), these practices are continuous and iterative, reflecting the ongoing use of the data and the dynamic nature of operational environments.

This review of administrative data should not be regarded as a one-off event, but is rather a process that requires repeated evaluation to understand the implications of changes and allow for the ongoing monitoring of the data quality.

Assessors will identify evidence showing the ongoing review of the administrative data by statistics producers.

## Quality Management Actions



## QA Matrix

The QA Matrix helps assessors and producers to determine the types of assurance and documentation required to inform users about the quality assurance arrangements for administrative data. This guidance can support a judgment about the suitability of the data and to identify examples of practices that meet the different levels of assurance.

Producers are responsible for judging the appropriate level of assurance. Producers should be able to explain the basis of their judgments of the chosen levels of assurance.

'No assurance' (A0) is not compliant with the *Code of Practice for Official Statistics*.

The need for investigation and documentation increases at each level of assurance from 'Basic' (A1) to 'Enhanced' (A2) to 'Comprehensive' (A3). It may be appropriate for the levels of assurance to vary among the four practice areas; for example, given specific circumstances it may be appropriate for 'Communication' to be Basic (A1), while 'Data Collection QA' be Enhanced (A2) and for both 'Operational Context' and 'Producer's QA' to be Comprehensive (A3).

Assessors will make an evaluation of what they regard as the appropriate level of assurance for the administrative data during an Assessment of official or National Statistics based on administrative data.

The Authority may decide that given the level of risk of quality issues and the public interest profile of the statistics that a higher level of assurance is appropriate than that judged by the statistics producer. In these cases, assurance levels A1 or A2 may be viewed as not compliant with the *Code*. The Assessment will identify the specific areas of practice that are required for compliance with the *Code*.

### Four practice areas associated with data quality

Operational context & admin data collection	Communication with data supply partners	QA principles, standards and checks by data suppliers	Producers' QA investigations & documentation
<ul style="list-style-type: none"> <li>• environment and processes for compiling the administrative data</li> <li>• factors which affect data quality and cause bias</li> <li>• safeguards which minimise the risks</li> <li>• role of performance measurements and targets; potential for distortive effects</li> </ul>	<ul style="list-style-type: none"> <li>• collaborative relationships with data collectors, suppliers, IT specialists, policy and operational officials</li> <li>• formal agreements detailing arrangements</li> <li>• regular engagement with collectors, suppliers and users</li> </ul>	<ul style="list-style-type: none"> <li>• data assurance arrangements in data collection and supply</li> <li>• quality information about the data from suppliers</li> <li>• role of operational inspection and internal/external audit in data assurance process</li> </ul>	<ul style="list-style-type: none"> <li>• QA checks carried out by statistics producer</li> <li>• quality indicators for input data and output statistics</li> <li>• strengths and limitations of the data in relation to use</li> <li>• explanation for users about the data quality and impact on the statistics</li> </ul>

## QA Matrix

Levels of assurance for four areas of practice related to quality assurance of administrative data regularly provided for producing official statistics (see Annex A to see the QA Matrix on one page):

### Practice area 1: Operational context & administrative data collection

Level of Assurance	Operational context & administrative data collection
<b>A0: No assurance</b>	<ul style="list-style-type: none"> <li>Operational context and administrative data collection by supplier not investigated, managed or documented</li> </ul>
<b>A1: Basic assurance</b> Statistical producer has reviewed and published a summary of the administrative data QA arrangements	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided users with an outline of the administrative data collection process,</li> <li>Illustrated the administrative data collection process and main stages,</li> <li>Outlined the operational context,</li> <li>Identified actions taken to minimise risks to quality,</li> <li>Identified and summarised the implications for accuracy and quality of data, including the impact of any changes in the context or collection arrangements</li> </ul>
<b>A2: Enhanced assurance</b> Statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided users with a fuller description of the operational context and administrative data collection arrangements, eg:                             <ul style="list-style-type: none"> <li>- a process map detailing the data collection processes,</li> <li>- explanations for classifications,</li> </ul> </li> <li>Identified and summarised potential sources of bias and error in administrative system,</li> <li>Identified and described safeguards taken to minimise risks to data quality,</li> <li>Provided a detailed description of the implications for accuracy and quality of data, including the impact of any changes in the context or collection arrangements</li> </ul>
<b>A3: Comprehensive assurance</b> Statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit, and published detailed documentation about the assurance and audit	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided users with a detailed description of the administrative system and operational context:                             <ul style="list-style-type: none"> <li>- explained why the data are collected, who by and how,</li> <li>- identified differences across areas in the collection and recording of the data,</li> <li>- identified issues for individual data items, such as whether objective or based on subjective recording, missing and/or imputed, poorly recorded,</li> </ul> </li> <li>Identified issues in design and definition of targets,</li> <li>Identified and described potential sources of bias and error in the administrative system,</li> <li>Identified and explained any safeguards used to minimise the risks to data quality,</li> <li>Provided a detailed and specific description of the implications for accuracy and quality of the data, including the impact of any changes in the context or collection arrangements</li> </ul>

**Practice area 2: Communication with data supply partners**

Level of Assurance	Communication with data supply partners
<p><b>A0: No assurance</b></p>	<ul style="list-style-type: none"> <li>• No communication</li> </ul>
<p><b>A1: Basic assurance</b> Statistical producer has reviewed and published a summary of the administrative data QA arrangements</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has outlined the data provision arrangements including:                             <ul style="list-style-type: none"> <li>- annual statement of needs,</li> <li>- timing and format of data supply,</li> <li>- coordination of data sign-off from data supplier,</li> </ul> </li> <li>• Fed back identified errors to data suppliers and recorded their response,</li> <li>• Sought the views of statistics users about the data and resolved any quality issues reported</li> </ul>
<p><b>A2: Enhanced assurance</b> Statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has agreed and documented:                             <ul style="list-style-type: none"> <li>- data requirements for statistical purposes,</li> <li>- legal basis for data supply,</li> <li>- data transfer process,</li> <li>- arrangements for data protection,</li> <li>- sign-off arrangements by data suppliers,</li> </ul> </li> <li>• Established an effective mode of communication with contacts (eg with data collector and supplier bodies, IT systems, operational/policy officials) to discuss the ongoing statistical needs in the data collection system and quality of supplied data,</li> <li>• Sought the views/experiences of statistics users and resolved any quality issues reported</li> </ul>
<p><b>A3: Comprehensive assurance</b> Statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit, and published detailed documentation about the assurance and audit</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has established/maintained collaborative relationships,</li> <li>• Has a written agreement specifying:                             <ul style="list-style-type: none"> <li>- roles and responsibilities,</li> <li>- legal basis for data supply,</li> <li>- data supply and transfer process,</li> <li>- security and confidentiality protection,</li> <li>- schedule for data provision,</li> <li>- content specification,</li> </ul> </li> <li>• Used a change management process,</li> <li>• Regularly communicated with the data collector and supplier bodies, IT systems, operational/policy officials eg newsletters, conferences, attending data supplier/IT system group meetings,</li> <li>• Regularly engaged statistics users, resolved any reported quality issues, and held user group conferences</li> </ul>

**Practice area 3: QA principles, standards and checks applied by data suppliers**

Level of Assurance	QA principles, standards and checks applied by data suppliers
<p><b>A0: No assurance</b></p>	<ul style="list-style-type: none"> <li>• No description of suppliers' QA procedures and standards</li> </ul>
<p><b>A1: Basic assurance</b> Statistical producer has reviewed and published a summary of the administrative data QA arrangements</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has knowledge of suppliers' QA checks and published a brief description,</li> <li>• Identified whether audits are conducted on the admin data (such as internal or operational audits, external audit such as by regulator),</li> <li>• Described the implications for the statistics</li> </ul>
<p><b>A2: Enhanced assurance</b> Statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has provided a fuller description of the main QA principles, quality indicators and checks used by the data suppliers,</li> <li>• Described the role of relevant information management or governance groups in data quality management,</li> <li>• Described the role of audit of the admin data within the collection and operational settings,</li> <li>• Described the implications for the statistics for the quality issues identified by data supply bodies and regulators</li> </ul>
<p><b>A3: Comprehensive assurance</b> Statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit, and published detailed documentation about the assurance and audit</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has described the data suppliers' principles, standards (quality indicators) and quality checks,</li> <li>• Reviewed quality reports for the received data (such as input quality indicators for data accuracy, coverage and completeness),</li> <li>• Identified and documented the findings of investigations and audits conducted on the admin data and associated targets (such as internal and operational audits, and external audits by regulators and professional bodies),</li> <li>• Described the implications for the statistics and determined whether the data continue to be satisfactory for official statistics purposes</li> </ul>



**Practice area 4: Producer’s QA investigations & documentation**

Level of Assurance	Producer's QA investigations & documentation
<p><b>A0: No assurance</b></p>	<ul style="list-style-type: none"> <li>• No description of own QA checks</li> </ul>
<p><b>A1: Basic assurance</b> Statistical producer has reviewed and published a summary of the administrative data QA arrangements</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has established regular QA checks on the received admin data,</li> <li>• Published a description of its own QA checks on the admin data,</li> <li>• Outlined the general approach and overall findings,</li> <li>• Identified the strengths and limitations of the admin data,</li> <li>• Explained the likely degree of risk to the quality of the admin data</li> </ul>
<p><b>A2: Enhanced assurance</b> Statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has provided a fuller description of its own QA checks on the admin data,</li> <li>• Detailed the general approach and findings for specific quality indicators,</li> <li>• Identified the strengths and limitations of the admin data,</li> <li>• Explained the likely degree of risk to the quality of the admin data</li> </ul>
<p><b>A3: Comprehensive assurance</b> Statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit, and published detailed documentation about the assurance and audit</p>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>• Producer has provided a detailed description of its own QA checks on the admin data (including validation, sense and consistency checks),</li> <li>• Given quantitative (and where appropriate qualitative) metrics for specific quality indicators (such as input, process and output quality metrics),</li> <li>• Undertaken comparisons with other relevant data sources (such as survey or other admin data),</li> <li>• Identified possible distortive effects of performance measurements and targets,</li> <li>• Identified the strengths and limitations of the admin data and any constraints on use for producing statistics,</li> <li>• Explained the likely degree of risk to the quality of the admin data</li> </ul>

## Risk/Profile Matrix

The critical judgment about the suitability of the administrative data for use in producing official statistics should be pragmatic and proportionate, made in the light of an evaluation of the likelihood of quality issues arising in the data that may affect the quality of the statistics, and of the nature of the public interest served by the statistics:

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

A1 to A3 = Levels of assurance in the QA Matrix

### Level of risk of data quality concerns

Use the following descriptions to help determine whether the administrative data should be regarded as low, medium or high risk of having quality concerns:

**Low risk** – the data may have a low risk of data quality concerns in situations in which there is a clear agreement about what data will be provided, when, how, and by whom; when there is a good appreciation of the context in which the data are collected, and the producer accepts that the quality standards being applied meet the statistical needs.

**Medium risk** – the data may be regarded as having a medium risk of data quality concerns when high risk factors have been moderated through the use of safeguards, for example, integrated financial audit and operational checks, and effective communication arrangements. It is also appropriate to consider the extent of the contribution of the administrative data to the official statistics, for example, in cases where the statistics are produced in combination with other data types, such as survey or census data.

**High risk** – the data may have a high risk of data quality issues when there are many different data collection bodies, intermediary data supplier bodies, and complex data collection processes with limited independent verification or oversight.

For example:

- The risk is lower when a supplier team is located within the same organisation as the statistics producer but increased when there are multiple data collection agencies and/or intermediary data suppliers.
- The risk is lower when operational processes include teams whose function it is to quality assure and verify the recorded data and where internal audit is an established

mechanism with the body. The risk is higher when no internal audit occurs and where limited external and independent challenge occurs to the data, such as, provided by regulator bodies.

- The risk is lower when clear agreements are in place and communication is regular and found to be effective by all partners in the data supply process. The risk of data quality issues is increased when roles and responsibilities are not clear and no formal agreement is in place for the supply of data.

### Public interest profile of the statistics

Statistical producers and assessors should use the following descriptions to help determine whether the administrative data should be regarded as having low, medium or high levels of public interest (representing the 'public good'):

**Low profile** – politically neutral subject; interest limited to niche user base, and limited media interest.

**Medium profile** – wider user and media interest, with moderate economic and/or political sensitivity.

**High profile** – economically important, reflected in market sensitivity; high political sensitivity, reflected by Select Committee hearings; substantial media coverage of policies and statistics; important public health issues; collection required by legislation.

### Top Tips

These five tips summarise the main pointers for statistical producers to develop a good understanding of the quality issues of administrative data:

Don't trust the safeguards	<ul style="list-style-type: none"> <li>• check if safeguards are functioning effectively</li> </ul>
Get involved	<ul style="list-style-type: none"> <li>• work and share with suppliers, such as through secondments and webinars, to develop a common understanding</li> </ul>
Raise a red flag	<ul style="list-style-type: none"> <li>• identify potential data quality concerns using input and output quality indicators and investigate anomalies</li> </ul>
See the big picture	<ul style="list-style-type: none"> <li>• identify what investigations and audits have been conducted and what they found</li> </ul>
Corroborate the evidence	<ul style="list-style-type: none"> <li>• confirm the levels and the trends shown by the stats derived from the admin data</li> </ul>

## Helpful Guidance

1. UKSA admin data webpage with case examples, talks from RSS event, Exposure Draft report and tool for non-statisticians:  
<http://www.statisticsauthority.gov.uk/assessment/monitoring/administrative-data-and-official-statistics/index.html>
2. NSO Guidance, Use of Administrative or Management Information:  
<https://gss.civilservice.gov.uk/blog/2014/05/interim-administrative-data-guidance/>
3. Using Administrative Data: Good Practice Guidance for Statisticians:  
<https://gss.civilservice.gov.uk/wp-content/uploads/2012/12/Interim-Admin-Data-guidance.pdf>
4. ESS network project on administrative data for business statistics – including Deliverable 2011/6.5, *Final list of quality indicators and associated guidance*, and Deliverable 2011/6.1, *Basic list of quality indicators* <http://www.cros-portal.eu/content/admindata-sga-3>
5. BLUE ETS Deliverable 4.2, Report on methods preferred for the quality indicators of administrative data sources: <http://www.blue-ets.istat.it/fileadmin/deliverables/Deliverable4.2.pdf>
6. ESS network Work Package 2a, Deliverables 2.2 and 2.4 – to check the suitability of an administrative data source before first use: <http://www.cros-portal.eu/content/admindata-sga-3>
7. US Federal Committee on Statistical Methodology, US Census Bureau data tool for checking a new administrative data source: [www.bls.gov/osmr/datatool.pdf](http://www.bls.gov/osmr/datatool.pdf)
8. ONS Quality Centre, Guidelines for Measuring Statistical Output Quality:  
<http://www.ons.gov.uk/ons/guide-method/method-quality/quality/guidelines-for-measuring-statistical-quality/index.html>

## Contacts

If you have any questions about this toolkit, please contact:

- [assessment@statistics.gsi.gov.uk](mailto:assessment@statistics.gsi.gov.uk)

Further advice about statistical practice can be sought from:

- the National Statistician's Good Practice Team (GPT):  
[goodpracticeteam@statistics.gsi.gov.uk](mailto:goodpracticeteam@statistics.gsi.gov.uk)
- ONS Quality Centre: [ons.qualitycentre@ons.gsi.gov.uk](mailto:ons.qualitycentre@ons.gsi.gov.uk)

Level of Assurance	Areas of practice related to quality assurance of administrative data regularly provided for producing official statistics			
	Operational context & administrative data collection	Communication with data supply partners	QA principles, standards and checks applied by data suppliers	Producer's QA investigations & documentation
<b>A0: No assurance</b>	<ul style="list-style-type: none"> <li>Context and data collection by supplier not investigated, managed or documented</li> </ul>	<ul style="list-style-type: none"> <li>No communication</li> </ul>	<ul style="list-style-type: none"> <li>No description of suppliers' QA procedures and standards</li> </ul>	<ul style="list-style-type: none"> <li>No description of own QA checks</li> </ul>
<b>A1: Basic assurance</b> Statistical producer has reviewed and published a summary of the administrative data QA arrangements	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided users with an outline of the administrative data collection process,</li> <li>Illustrated the administrative data collection process and main stages,</li> <li>Outlined the operational context,</li> <li>Identified actions taken to minimise risks to quality,</li> <li>Identified and summarised the implications for accuracy and quality of data, including the impact of any changes in the context or data collection</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has outlined the data provision arrangements including:                             <ul style="list-style-type: none"> <li>annual statement of needs,</li> <li>timing and format of data supply,</li> <li>coordination of data sign-off from supplier,</li> </ul> </li> <li>Fed back identified errors to data suppliers and recorded their response,</li> <li>Sought the views of statistics users about the data and resolved any quality issues reported</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has knowledge of suppliers' QA checks and published a brief description,</li> <li>Identified whether audits are conducted on the admin data (such as internal or operational audits, external audit such as by regulator),</li> <li>Described the implications for the statistics</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has established regular QA checks on the received admin data,</li> <li>Published a description of its own QA checks on the admin data,</li> <li>Outlined general approach and findings,</li> <li>Identified the strengths and limitations of the admin data,</li> <li>Explained the likely degree of risk to the quality of the admin data</li> </ul>
<b>A2: Enhanced assurance</b> Statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided users with a fuller description of the operational context and administrative data collection arrangements, eg:                             <ul style="list-style-type: none"> <li>a process map detailing data collection processes,</li> <li>explanations for classifications,</li> </ul> </li> <li>Identified and summarised potential sources of bias and error in administrative system,</li> <li>Identified and described safeguards taken to minimise risks to data quality,</li> <li>Provided a detailed description of the implications for accuracy and quality of data, including the impact of any changes in the context or collection arrangements</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has agreed and documented:                             <ul style="list-style-type: none"> <li>data requirements for statistical purposes,</li> <li>legal basis for data supply,</li> <li>data transfer process,</li> <li>arrangements for data protection,</li> <li>sign-off arrangements by data suppliers,</li> </ul> </li> <li>Established an effective mode of communication with contacts (eg with data collector and supplier bodies, IT systems, operational/policy officials) to discuss the ongoing statistical needs in the data collection system and quality of supplied data,</li> <li>Sought the views/experiences of statistics users and resolved any quality issues reported</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided a fuller description of the main QA principles, quality indicators and checks used by the data suppliers,</li> <li>Described the role of relevant information management or governance groups in data quality management,</li> <li>Described the role of audit of the admin data within the collection and operational settings,</li> <li>Described the implications for the statistics for the quality issues identified by data supply bodies and regulators</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided a fuller description of its own QA checks on the admin data,</li> <li>Detailed the general approach and findings for specific quality indicators,</li> <li>Identified the strengths and limitations of the admin data,</li> <li>Explained the likely degree of risk to the quality of the admin data</li> </ul>
<b>A3: Comprehensive assurance</b> Statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit, and published detailed documentation about the assurance and audit	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided users with a detailed description of the administrative system and operational context:                             <ul style="list-style-type: none"> <li>explained why the data are collected, who by and how,</li> <li>identified differences across areas in the collection and recording of the data,</li> <li>identified issues for individual data items, such as whether objective or based on subjective recording, missing and/or imputed, poorly recorded,</li> </ul> </li> <li>Identified issues in design and definition of performance measurements and targets,</li> <li>Identified and described potential sources of bias and error in the administrative system,</li> <li>Identified and explained any safeguards used to minimise the risks to data quality,</li> <li>Provided detailed description of the implications for a quality of the data, including the impact of any changes in the context or data collection</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has established/maintained collaborative relationships,</li> <li>Has a written agreement specifying:                             <ul style="list-style-type: none"> <li>roles and responsibilities,</li> <li>legal basis for data supply,</li> <li>data supply and transfer process,</li> <li>security and confidentiality protection,</li> <li>schedule for data provision,</li> <li>content specification,</li> </ul> </li> <li>Used a change management process,</li> <li>Regularly communicated with the data collector and supplier bodies, IT systems, operational/policy officials eg newsletters, conferences, attending data supplier/IT system group meetings,</li> <li>Regularly engaged statistics users, resolved any reported quality issues, and held user group conferences</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has described the data suppliers' principles, standards (quality indicators) and quality checks,</li> <li>Reviewed quality reports for the received data (such as input quality indicators for data accuracy, coverage and completeness),</li> <li>Identified and documented the findings of investigations and audits conducted on the admin data and associated targets (such as internal and operational audits, and external audits by regulators and professional bodies),</li> <li>Described the implications for the statistics and determined whether the data continue to be satisfactory for official statistics purposes</li> </ul>	<p><i>Consider the following types of activities:</i></p> <ul style="list-style-type: none"> <li>Producer has provided a detailed description of its own QA checks on the admin data (including validation, sense and consistency checks),</li> <li>Given quantitative (and where appropriate qualitative) metrics for specific quality indicators (such as input, process and output quality metrics),</li> <li>Undertaken comparisons with other relevant data sources (such as survey or other admin data),</li> <li>Identified possible distortive effects of performance measurements and targets,</li> <li>Identified the strengths and limitations of the admin data and any constraints on use for producing statistics,</li> <li>Explained the likely degree of risk to the quality of the admin data</li> </ul>