



**Quality Assurance of Administrative Data:
Frequently Asked Questions for Statistical Producers**

May 2016

V1.1

Contact:

Penny Babb

UK Statistics Authority

penny.babb@statistics.gov.uk

List of Questions

1. What is the definition of 'admin data'?	3
2. Do I need to bother with QAAD if I don't think my data are admin data?	3
3. Who are data supply partners?	3
4. Multiple OS producers – who has what responsibility?	3
5. Who does what to QA the data when there are different teams in a producer body involved in collating the data and publishing the output?	4
6. Where should we publish the QAAD documentation? Does it need to be in a separate document?	4
7. How much of the quality information needs to be documented or published for users?	5
8. Why do I need to bother as my users aren't interested in reading quality information?	6
9. How can I find out how other producers have responded to QAAD?	6
10. How do I work out the level of public interest of the statistics?	6
11. How do I decide on the level of data quality concern?	7
12. What about when there are many data sources used for the same statistics – does it mean all of them need the same level of assurance?	8
13. How do the Quality Management Actions fit into the QA Matrix?	9
14. How can I ask the suppliers whether they QA the data without annoying them or causing offence? We rely on their good will.	9
15. We are responding to QAAD because of an Assessment requirement but we can't finish the work within the Assessment deadline - what can we do to get more time?	10
16. How does the Authority ensure it applies the guidance consistently across all the assessments when assessing official statistics?	10

Change since v1.0:

1. Hyperlinks to footnotes 4 and 5 have been updated to reflect website changes

Quality Assurance of Administrative Data (QAAD):

FAQs for statistical producers

This Q&A document supports statistical producers in applying the Statistics Authority's QAAD guidance¹ when quality assuring data for use in producing official statistics.

1. What is the definition of 'admin data'?

Admin data are data that are collected for administrative or operational purposes, with statistical use being a secondary purpose.

2. Do I need to bother with QAAD if I don't think my data are admin data?

The *Code of Practice for Official Statistics* requires all OS to be quality assured, for their quality to be sufficient for the use made of the statistics and explained to users. QAAD represents fundamental good practice when considering the quality of your data – it is worth applying the guidance irrespective of the nature of the data source, to ensure that you have sufficient insight into any data quality issues and their impact on your statistics.

3. Who are data supply partners?

Data supply partners are any organisations involved in data collection and supply, such as the data collection organisation, IT systems teams, MI managers, policy and operational officials. They can include contractors, as well as public sector bodies. There can be just one organisation and there can be multiple bodies. There can be intermediaries, such as when a contractor processes data from the data collection bodies prior to passing the data to the statistical producer body.

The more complex is the relationship or process, the greater the potential risk for data quality issues.

4. Multiple OS producers – who has what responsibility?

For many official statistics, one statistical producer body may collect the admin data and pass it on to another OS producer for preparing and publishing an additional ('secondary') set of OS. The primary statistical producer, i.e. the producer output team producing the first/main set of official statistics from the source, has the primary responsibility for liaising with data supply partners. The primary output producer body

¹ <https://www.statisticsauthority.gov.uk/monitoring-and-assessment/monitoring/administrative-data-and-official-statistics/>

should provide summary information to secondary OS producer bodies, setting out the data quality issues and the nature of the assurance arrangements.

The secondary OS producer bodies, though, have an essential responsibility to consider the nature of the data source and its ongoing suitability for producing their OS. In particular, they should consider whether there are specific issues relevant to their statistics that require further investigation or assurance. For example, it may be the case that the secondary output is based on variables not used in producing the primary output statistics. In such cases, it may be appropriate for the secondary statistical producer body to seek further assurance from the primary statistical output producer or to contact the data supply partners directly.

Essential: It is important that each statistical output team can explain to users why they are satisfied that the data are sufficiently robust for their use in producing the official statistics and set out the strengths and limitations of the statistics in relation to use.

5. Who does what to QA the data when there are different teams in a producer body involved in collating the data and publishing the output?

The data receiving team should: establish a good understanding about the data source; build strong ties with the supply partners, including an understanding of the QA applied by the partners; and undertake their own quality assurance of the data received from the suppliers. The data receiving team should provide information to the statistical output team about the quality of the data and its assurance.

The statistical output team should ensure that it understands the data collection process and any issues that may impact the statistics and their use. It will conduct its own QA of the data, such as sense checks, consistency checks over time, reviewing quality indicators from the suppliers, and comparisons with other relevant sources. It may rely on gaining insight about the source of the data and the collection process from the producer body's data receiving team. On occasion, though, the output team may need to directly contact data supply partners, to understand particular issues that impact the secondary statistics.

Communication is at the heart of ensuring a common understanding – all have a part to play, but the output team is responsible for considering the implications for the statistics and explaining to users about any issues that may have a substantive effect on the statistics.

6. Where should we publish the QAAD documentation? Does it need to be in a separate document?

The Authority is not prescriptive about where to publish explanatory information about the quality assurance arrangements and data quality issues impacting the statistics. The statistical producer team should consider the needs of users and how to ensure that the

information is accessible and clear. It is helpful to provide a summary statement with the statistics setting out that you are confident that the data are sufficiently robust to produce the official statistics and why.

Producers have adopted a variety of approaches to publishing QA information – this catalogue [add link] on the GSS website provides a list of many of the published documents.

The Authority has produced some case examples² outlining some areas of good practice in assuring the data and describing quality issues.

7. How much of the quality information needs to be documented or published for users?

The QAAD guidance emphasises three levels of assurance – fuller documentation is required with the higher levels than for the lowest level. The amount of explanatory material depends on the nature of the specific issues and the needs of users, to support their appropriate use of the statistics.

It is important to provide some summary information alongside the statistics but output producers can provide further supporting information in separate documents (with clear signposting). It may be helpful to provide a summary in a quality report in the appropriate parts of the ESS framework. It is not always necessary to produce a separate QA document but in some circumstances that can be a helpful choice.

For example, HSCIC's NHS Outcomes Framework³ provides information about the QA of admin data in its background quality notes for the each indicator. In contrast, NI Department of the Environment⁴ published a QAAD assessment report⁵ after its investigations into the QA arrangements of the variety of admin data sources used to produce driver and vehicle statistics. The Home Office summarised the data quality assurance and quality issues for its terrorism statistics within its existing user guide⁶. The Department for Education published an overview of the assurance of the school-level examination statistics on its website⁷ for parents, plus additional detailed guidance for expert users in its quality reports⁸.

² <https://www.statisticsauthority.gov.uk/monitoring-and-assessment/monitoring/administrative-data-and-official-statistics/quality-assurance-of-administrative-data-case-examples/>

³

<http://www.hscic.gov.uk/searchcatalogue?q=title%3A%22nhs+outcomes+frameworks+indicators%22&area=&size=10&sort=Relevance>

⁴ <https://www.infrastructure-ni.gov.uk/topics/statistics-and-research/driver-vehicle-agency-statistics>

⁵ <https://www.infrastructure-ni.gov.uk/publications/driver-and-vehicle-statistics-northern-ireland-administrative-data-quality-assessment>

⁶ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/513730/user-guide-police-powers-terror.pdf

⁷ <http://www.education.gov.uk/schools/performance/aboutadmindata.html>

⁸

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/483204/SFR47_2015_Qualitvandmethodology.pdf

8. Why do I need to bother as my users aren't interested in reading quality information?

The Authority's primary concern in producing the QAAD guidance was to encourage statistical producers to develop a probing, challenging approach to quality assuring admin data and achieve sufficient understanding of the issues that impact their official statistics. In doing this, it is important that producers explain to users why they are confident that the data are robust for their use. These investigations may highlight concerns that are relevant to users, to help guide their appropriate use.

These are not new requirements under the *Code of Practice* but the QAAD guidance helps draw attention to a wide range of statistical practices that can be used when quality assuring data. The four practice areas emphasise that QA is beyond a producer's own QA checks. Appropriate assurance includes understanding the reason for the data collection and the nature of the process. It also includes the relationships developed with the various data supply partners – from the initial collection and operational or policy development, through the IT system management, to the governance arrangements and communication with users. Appropriate assurance also requires an understanding of the nature of any QA carried out on the data before it reaches the statistical output team. Where analysts consider the admin data in each of these ways, they will gain a rounded picture of the data and any weaknesses that may impact the statistics.

9. How can I find out how other producers have responded to QAAD?

The Authority has produced some case examples on its website to illustrate some areas of good practice demonstrated by statistical producers in responding to QAAD guidance. It has also produced a catalogue [add link] on the GSS website with links to published QA documents from OS producers.

10. How do I work out the level of public interest of the statistics?

The first thing to note is that it is the public interest or value of the official statistics, as opposed to the individual data source. There are some questions that you can ask yourself when deciding whether the public interest profile is lower, medium or higher:

- What use is made of the statistics?
- What decisions do they impact? (eg spending by government)
- What is the reputation risk attached to the statistics?
- How broad is the user base?
- Is there political interest in the statistics? (eg government commitments)
- Are they required by legislation?
- Are the statistics used to hold the government to account?

Level of public interest or value	Example criteria
Lower	Always likely to be politically neutral subject Interest 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 eg by government Highly political sensitivity eg reflected by Select Committee hearings Substantial media coverage of policies and statistics Substantial public health concern

11. How do I decide on the level of data quality concern?

Consider the characteristics of the data collection and the reasons for the data:

- Where are the data from?
- Why are they collected?
- How are the data entered? eg manually vs automatically
- How many organisations are involved? eg one vs many
- Why does it matter to my users if the data I am using is of poor quality?
- Have the data collection or processing systems changed or are they changing?
- Are the data supplied under contract?
- Does the contract contain key quality indicators or standards?
- Have there been policy changes that have changed the data collection requirements?
- How many errors, delays, incomplete or resubmission of data have I had?
- Do I pay for the data?

Level of data quality concern	Example criteria
Lower	<ul style="list-style-type: none"> 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	<ul style="list-style-type: none"> 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 eg by regulators Multiple data providers offset by use of quality indicators
Higher	<ul style="list-style-type: none"> 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

12. What about when there are many data sources used for the same statistics – does it mean all of them need the same level of assurance?

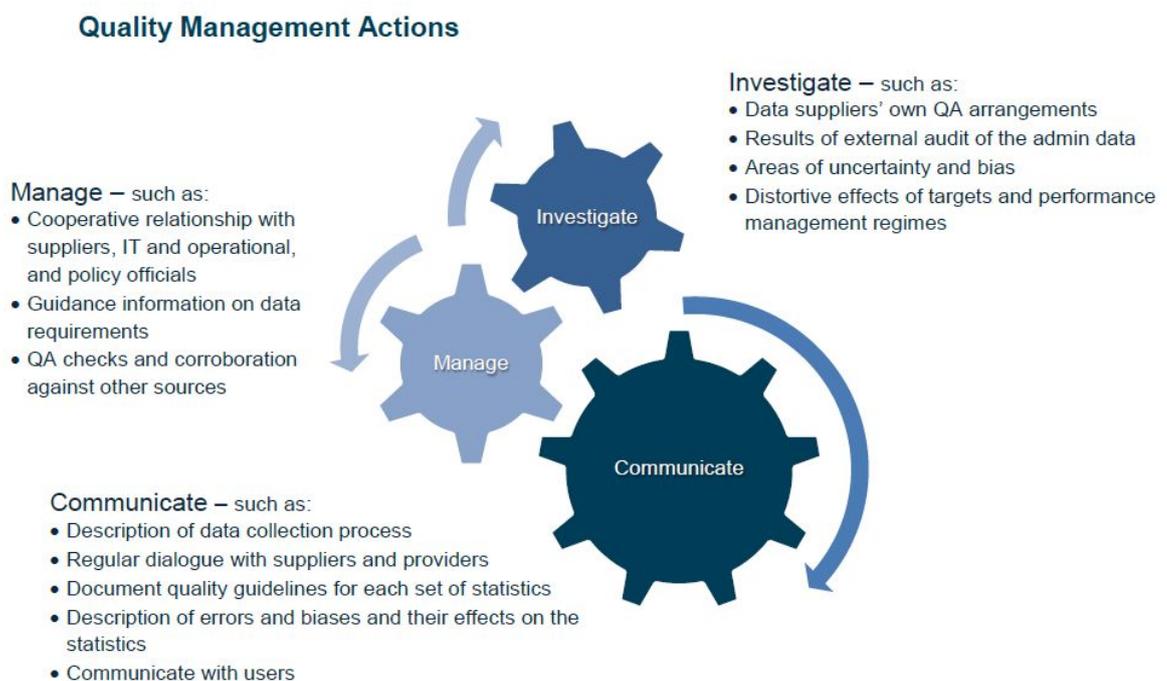
While the public interest profile reflects the statistics, the risk of data quality concerns reflects the individual admin data source. When producing statistics based on a number

of data sources, first consider the risk of data quality concerns for each source and then determine the impact of the individual source on the statistics. If an individual data source has a nominal impact on the statistics, it would be appropriate to reduce the level of data quality concern of that source, and hence the overall level of assurance.

In considering the data quality concerns, it may be helpful to contrast the various sources, to see which cause you greatest concern. It is also worth identifying any ways that the method of producing the statistics overcomes (or mitigates) the weaknesses in the data. Always ask yourself what evidence there is to show that known safeguards are working correctly.

13. How do the Quality Management Actions fit into the QA Matrix?

In developing techniques for considering the quality of admin data, it is essential to build in ongoing ways of testing and checking – to avoid assumptions that all is well with the data into the future. The QMA is a reminder of areas to consider when developing these ongoing statistical practices in statistical teams.



14. How can I ask the suppliers whether they QA the data without annoying them or causing offence? We rely on their good will.

Developing good relationships and effective communication are at the heart of gaining a good understanding of data issues. When contacting suppliers to find out about their assurance approach, it may be worth emphasising that it is the Statistics Authority's recommendation that you gain a fuller understanding of the ways that the data are collected and why, and what steps the suppliers undertake to verify and assure the accuracy of the data. Their information will provide greater reassurance about the

suitability of the data. It may also give a greater insight into issues affecting the statistics.

Meeting with suppliers can help them to get to know you and build a better understanding of their situation. You might find that it is possible to join data governance or information management groups.

15. We are responding to QAAD because of an Assessment requirement but we can't finish the work within the Assessment deadline - what can we do to get more time?

The Assessors will discuss the timing of addressing any requirements with you during the final stages of the Assessment and when you expect to be able to meet the requirements. It may be that you need more time to meet requirements related to the QAAD guidance – the Authority can allow for this in the Assessment report deadline. Be sure to discuss any concerns with the Assessment team.

Should you experience difficulties or delays after the publication of the Assessment report, just discuss the issues with the Assessment team and they can consider these in their follow up. It may be that your team would benefit from attending a QAAD workshop – let us know as we can provide QAAD workshops.

16. How does the Authority ensure it applies the guidance consistently across all the assessments when assessing official statistics?

The Assessment process has a number of checks built in, to challenge and review the evidence and judgments made by the Assessment teams. These include a QA stage involving a review of the evidence by senior managers, a peer review by another assessor not otherwise involved in the assessment, and a review of the draft assessment report by senior managers, including the DG for Regulation. These stages ensure consistency in the judgments made by the individual assessors.

The GSS Good Practice Team receives each draft Assessment report. They look for consistency in the application of the *Code* and with other assessments.