Part 2: Learning from current quality assurance and audit practice

Introduction

2.1 In seeking to establish a standard for appropriate audit arrangements, we investigated the practices currently employed in producing six sets of official statistics. This section presents some core learning from our review of these case studies, alongside areas where we noted either existing practices, or opportunities for the development of such practices.

2.2 The reviews of these statistics are presented as case studies in Annex C:
- Office for National Statistics' police recorded crime statistics
- Health and Social Care Information Centre's (HSCIC) social services activity and expenditure statistics
- HSCIC's hospital episode statistics
- Information Services Division's prescription statistics in Scotland
- Department for Work and Pension's Work Programme statistics
- Health and Safety Executive's (HSE) injury statistics

2.3 The case studies describe the quality assurance and audit arrangements for the selected official statistics and provide helpful examples for statistical producers to learn from others' experiences. We reviewed documentation provided for the Assessment of compliance with the *Code*, as well as published material associated with the statistics. We also drew on information from discussions with the statistical producer teams. We identified additional material about each of the administrative data sources and other relevant investigations or associated documentation such as reports published by NAO or regulator bodies.

2.4 We selected these case studies in order to include data from a range of different types of administrative sources and circumstances, such as multiple suppliers in local authorities and health trusts, payment systems with integrated financial audit, and an executive agency with combined responsibilities for data collection and statistics production. ONS's police recorded crime statistics are an important example as they are from an underlying administrative system with a number of established safeguards, but also one which has received considerable attention due to public concerns about the accuracy of the data and the limited assurance provided by the safeguards.

2.5 We now describe some lessons revealed by the case studies that illustrate some ways that the practices have been used to better understand the quality of statistics derived from administrative data.

Lesson 1: Don't trust the safeguards

2.6 It is clear from the police recorded crime case study (C1) that a formalised process for the management of administrative data and its assurance may not be enough in itself. The environment in which the organisations are operating may also play a role in influencing the handling of information and the wider context needs to be understood by statistical producers. For example, the producers should identify the potential distortive effect of targets and performance management regimes on the behaviour of those compiling the administrative information. These pressures should be addressed by the data supplier bodies with further appropriate safeguards taken. Statistical producers need to identify the associated weaknesses of the data and determine whether the safeguards are functioning effectively, and what more might be done to raise the level of assurance. They should ensure that they communicate clearly the implications for the statistics in relation to their use.

Lesson 2: Get involved

2.7 HSCIC has established strong ties with local authorities (LAs) in England with responsibility for adult social care services (case study C2). It has arranged secondments of LA staff to work in HSCIC. This opportunity ensured that a good understanding of the operational context underpinned the development of the new data collections by HSCIC, after a thorough review of adult social care data and users' needs. The importance of establishing clear agreements and relationships between statistical producers and data suppliers is also highlighted in the National Statistician's guidance on the use of administrative data.

Lesson 3: Raise a red flag

2.8 HSCIC's Hospital Episode Statistics case study (C3) illustrates the challenge of having an extremely large number of data collectors and suppliers – there are around 700 health trusts. In order to make the investigation of data quality concerns manageable, HSCIC draws on its effective engagement with supplier representative bodies and information governance groups, to identify potential data quality concerns. It also uses its own quality assurance of the supplied data to identify suspected issues or 'red flags'. For example, HSCIC reviews data for all the health trusts to identify whether any data suppliers are outliers that require further investigation. Audit information about the practices of the individual trusts may provide additional evidence for statistical producers when seeking to determine the adequacy of operational arrangements. The compilation of performance indicators across the supplier organisations, for example, for benchmarking, could also be used by statistical producers to further corroborate the effectiveness of the organisational arrangements in the supplier bodies. These types of evidence present an opportunity for further investigation to

support the statistical team's judgment about the suitability of the organisations' practices in producing and supplying the administrative data. HSCIC recognises the importance of establishing the quality of the data to meet operational or clinical needs, as well as for the secondary use of producing official statistics. Sharing this type of information across relevant statistical teams would help the statistical producer gain maximise benefit from the investigations.

Lesson 4: See the big picture

2.9   Central to judging the appropriateness of audit arrangements is to identify what specific audits and investigations have been conducted, how frequently and by whom. The case studies on ISD's prescription statistics (C4) and DWP's Work Programme statistics (C5) are based on payment information, and financial audit is an inherent part of the checks conducted on the data. Both case studies also include examples of process maps, with the ISD example indicating the points at which checks are made. The DWP case study demonstrates the detailed checks made to ascertain the suitability of the administrative source for use in producing official statistics. It also places the financial audit within the context of the quality assurance checks carried out by the statistical team.

Lesson 5: Corroborate the evidence

2.10   HSE (C6) has identified an under-recording of injuries in its non-fatal injury data, and so commissions questions on the Labour Force Survey to measure self-reported injury. This example highlights the fact that weaknesses in some administrative data sources cannot necessarily be overcome. Statisticians may need to seek alternative information to corroborate the administrative data. HSE uses different types of checks to quality assure the injury data: thorough checks by the HSE Inspectors, built-in system checks when employers notify HSE of an injury, internal audit of the system itself and its subsequent improvement, and the use of sample audit to check the completeness of the injury data. The audit of sampled data was found to be a particularly important source of evidence, to support the use of the administrative data, both for DWP and ISD, as well as HSE.

Conclusion

2.11   We identified some areas of good practice in the assurance and auditing of administrative data, as well as areas where there are opportunities for statistical producers to provide further information about quality assurance and audit arrangements – see Table 2.1 for a summary. We found that there is often a great deal of existing information available that is not accessed or used by the statistical producers to assess the quality of the data. A number of the practices identified in the case studies are specifically mentioned in the National

Statistician's existing guidance[8] about the use of administrative data – this guide emphasises the need to investigate the data source, work closely with data suppliers, and understand possible causes of error through validation and triangulation with other sources. Determining the suitability of the data should not be considered a one-off judgment but an ongoing process of challenge. Concerns identified through statistical producers' own quality assurance checks provide an opportunity to identify weaknesses or risks within the data supply process. New issues can emerge as operational changes occur to the administrative systems, as well as the implementation of classification and coding changes.

2.12   Each of the producers has developed detailed checks on the data received from the suppliers. Having identified anomalies, they request that the data suppliers investigate further the issues and possibly resubmit the data. Very often, though, the investigations focus solely on the internal validity of the data. We found that statistical producers tend not to use the quality issues they have found as prompts to probe the data suppliers' audit arrangements, or identify systemic or organisational issues that contribute to poor data.

2.13   In Part 3 we extend the areas of good practice identified in the six case studies to suggest a conceptual framework for evaluating the assurance and audit arrangements of administrative data in ways that are consistent with the *Code*. We then propose some tools to help statistical producers make this critical evaluation of the administrative data. These require the producer to first gauge the level of concern about data quality and the importance of the statistics derived from the administrative data. A maturity matrix (the '**Quality Assurance (QA) Matrix**') then guides the producer in determining the scale of investigation and documentation that are consistent with the data quality concerns and the nature of the use of the statistics.

---

[8] https://gss.civilservice.gov.uk/wp-content/uploads/2012/12/Interim-Admin-Data-guidance.pdf

Table 2.1: Examples of audit and assurance practices arising from the six case studies

Practice	Case study practice example
Determine the capacity and capability of the data supplier organisation(s) to provide accurate and reliable data:	
• Gather information about assurance and audit arrangements from external data suppliers	ISD's prescriptions statistics
• Develop a clear process map	ISD's prescriptions statistics
• Identify audit and assurance issues at institutional and operational levels	HSE's injury statistics
• Identify the potential impact of targets and performance management regimes	ONS's police recorded crime statistics
• Identify the results of internal audit of the systems and processes for administrative data	DWP's Work Programme statistics
• Identify the frequency and nature of external oversight for high profile sets of statistics	ONS's police recorded crime statistics
Collate organisational/operational information in a form that is readily accessible by statisticians in making judgments about the suitability and accuracy of the administrative data	HSCIC's HES statistics
Develop co-operative arrangements with data supplier organisation(s)	HSCIC's social services statistics
Identify how clearly the roles of those who record, quality assure and sign off the data for delivery are specified and understood	ONS's police recorded crime statistics
Establish whether a Memorandum of Understanding or a Service Level Agreement exists between the statistical producers and data suppliers:	
• Identify how clearly the roles and responsibilities of those involved in the process are specified	ONS's police recorded crime statistics
• Determine whether the key points from this agreement have been published	ONS's police recorded crime statistics
Fully document audit arrangements and identify the implications for the statistics:	
• Explain each stage of data collection, processing and quality assurance, demonstrating who is responsible at each stage and the checks that they carry out on the data	HSCIC's HES statistics
• Identify the potential risks to the accuracy of the data and the safeguards that are in place to minimise the risk	DWP's Work Programme statistics
Identify alternative data and information sources e.g. from surveys – to verify accuracy of the data	HSE's injury statistics