

Annex C: Case studies illustrating the use of information about the audit of administrative data by producers of official statistics

C.1 ONS's Police Recorded Crime Statistics²³

Background to the data

Until 2012 police recorded crime (PRC) statistics were collated and published by the Home Office. In April 2012, responsibility for the production of these statistics moved to the Office for National Statistics (ONS). ONS also took over responsibility for the Crime Survey in England and Wales, with TNS-BMRB commissioned to conduct the survey. The Home Office retains responsibility for policy making in this area. It continues to collate and quality assure the PRC data prior to sending them to ONS for the production of the crime official statistics. There is a (unpublished) Memorandum of Understanding between ONS and the Home Office which outlines the roles and responsibilities of each department in the production of statistics about crime. ONS publishes a quarterly report which presents statistics from both PRC and the Crime Survey and it also publishes topic based reports throughout the year.

In January 2014, the UK Statistics Authority found that there was an accumulation of evidence to suggest that the underlying data on crimes recorded by the police may not be reliable and it removed the designation of National Statistics from the recorded crime statistics. It was retained by the Crime Survey for England and Wales.

ONS's recorded crime statistics are used by the police to monitor trends and by the Home Office to design and monitor policies and strategies. Other government departments use specific aspects of the statistics to monitor their own policies – for example Ministry of Justice uses analyses focusing on perceptions of crime and criminal justice to inform the development of criminal justice system reform policy. Police and Crime Commissioners use the statistics to make comparisons with similar forces and regional and national averages, and to monitor local targets. Local authorities use the statistics to gain a regional picture of crime rates and to monitor the impact of policy reforms on council services, service users and the broader community. The statistics are regularly used by ONS and the Home Office to answer Parliamentary Questions about crime. The statistics are widely reported in the media, reflecting the public interest in crime and the criminal justice system. Academics use the statistics as part of a wide range of criminal justice research and for teaching purposes. Voluntary organisations use them to assess the risks of victimisation across different groups.

PRC data are supplied to the Home Office every month by the 43 police forces in England and Wales, plus the British Transport Police. Data are submitted either via aggregated returns (Excel spreadsheets) or the new HO Data Hub which provides

²³ <http://www.ons.gov.uk/ons/rel/crime-stats/crime-statistics/index.html>

record level data. The police are required to record crime in line with the National Crime Recording Standard (NCRS) and the Home Office Counting Rules for Recorded Crime (HOCR). All those who record crimes in police forces should be trained in the application of these standards.

Each police force has its own force crime registrar, who is responsible for overseeing compliance with HOCR and NCRS, and who is the final arbiter for whether a force should record a crime or make a 'no-crime' decision (to reverse the decision that an incident was a crime). Crimes can be reported in a variety of ways, for example to a call centre, to the police in person on the street, or by a third party and are recorded onto individual police force systems. The facts recorded about each incident allow the police to assess the matter reported to them and determine whether a crime has been committed.

Quality assurance

The Home Office Statistics Unit's Police Data Collection Section ensures that each force has submitted data and carries out basic quality assurance checks, such as comparing the current month's return with previous ones, identifying outliers or inconsistencies in the data. It raises queries with individual forces if there are revisions outside a certain tolerance level or if there are missing data or obvious errors found. The forces are then asked to check their data and resubmit them if necessary. The PRC data are then sent to the Home Office Crime Statistics team who carry out more validation checks and raise further queries with forces about data that appear inconsistent. A process map showing the stages of data quality assurance throughout the data cycle would be useful.

The Home Office supplies the PRC data to ONS who then carry out some further independent consistency checks, and examines the consistency with the crime survey data for equivalent offences to inform users about the relationship between the two series.

Audit

Police forces have internal audit procedures and the reports are generally discussed at each force's audit committee. After the launch of the NCRS in 2002, and up until 2007 the Audit Commission, commissioned by the Home Office, audited a sample of incident reports to check if crimes had been recorded correctly and published the results²⁴. For a number of years following completion of the Audit Commission's 2007 work, there was a lack of external scrutiny of recorded crime data.

Her Majesty's Inspectorate of Constabulary (HMIC) has the remit to carry out independent inspections of policing, including how crimes are recorded. Until recently HMIC had to be commissioned to carry out inspections, this is no longer the case. HMIC has carried out a number of recent reviews:

²⁴ <http://archive.audit-commission.gov.uk/auditcommission/nationalstudies/communitysafety/Pages/policedataquality0607.aspx.html>

- In 2012 HMIC published a review²⁵ of the quality of crime and incident data recorded by all 43 police forces in England and Wales plus the British Transport Police, and the arrangements in place to ensure that standards are maintained and improved. For the purpose of this review, a key finding was that HMIC ‘found limited evidence of forces directly assessing whether their own crime quality audits provided confidence that their crime figures gave an accurate account of their performance. Few forces compare crime audits with crime performance in any meaningful way’.
- In 2013 the Police and Crime Commissioner for Kent commissioned HMIC to conduct an inspection into crime recording in Kent Police. The report²⁶ concluded that ‘appreciably more needed to be done before the people of Kent could be confident that the crime and resolution figures published by the force were as accurate as they should be’.
- HMIC is currently carrying out an inspection of the integrity of crime reporting in England and Wales. It published an interim report²⁷ in April 2014 which highlighted some concerns with crime recording at the forces covered by the interim report. The full report is due to be published in autumn 2014.

Findings

- A lack of external scrutiny over a number of years for this high profile set of statistics has contributed to concerns about the underlying data, as noted in our earlier Monitoring Review ‘Overcoming Barriers to Trust in Crime Statistics’²⁸. The Public Administration Select Committee’s (PASC) recent report²⁹ welcomed HMIC’s decision to undertake a crime data integrity inspection in 2014.
- There is no clear outline of the process of data collection and quality assurance, in police forces, the Home Office or ONS. The statisticians could produce a process map to show the responsibilities of each party at each step in the process, and the stages of the quality assurance processes, identifying the areas of potential risk to the quality and accuracy of the data and the safeguards that are in place to minimise these risks. For example, a clear risk is at the start of the process if crimes are not recorded when they should be.
- Force Crime Registrars (FCRs) are of differing levels of seniority in police forces across England and Wales. PASC’s recent report noted that it is ‘essential that the Force Crime Registrar has not only had the requisite training but the necessary authority within the force to do their job’. FCRs should have clear backing from Chief Constables to record crimes with integrity.

²⁵ <http://www.hmic.gov.uk/media/review-police-crime-incident-reports-20120125.pdf>

²⁶ <http://www.hmic.gov.uk/publication/crime-recording-in-kent/>

²⁷ <http://www.hmic.gov.uk/publication/crime-recording-a-matter-of-fact-interim-report/>

²⁸ <http://www.statisticsauthority.gov.uk/reports---correspondence/reports/overcoming-barriers-to-trust-in-crime-statistics--england-and-wales.pdf>

²⁹ <http://www.publications.parliament.uk/pa/cm201314/cmselect/cmpubadm/760/760.pdf>

- ONS published a methodological paper³⁰ analysing the variation in crime trends between the Crime Survey for England and Wales and the police recorded crime data. This paper highlighted the growing divergence between the two sources since the cessation of external scrutiny in 2007.

Wider lessons

- High profile sets of statistics should be the subject of regular external scrutiny
 - Statisticians should make clear the level of scrutiny that they consider is necessary for each set of statistics
- Consider the roles of those who record, quality assure and sign off the data for publication
 - The statisticians should satisfy themselves that the data are managed by those who are sufficiently independent from reliance on targets or in the position of target setting, or whose performance will be judged on the basis of the data, within the data provider organisation
- Establish a Memorandum of Understanding or a Service Level Agreement between the statistical producer and data supplier bodies
 - This should clearly state the roles and responsibilities of those involved in the process
 - The key points from this agreement should be published
- Develop a clear process map
 - 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
 - This is especially important where data collections are complex and where several key stakeholders are involved in compiling the statistics.
- Identify the potential risks to the accuracy of the data and assess the safeguards that are in place to minimise the risk
- Carry out analysis of other key data sources which are used to corroborate the underlying data.

³⁰ <http://www.ons.gov.uk/ons/guide-method/method-quality/specific/crime-statistics-methodology/methodological-note--analysis-of-variation-in-crime-trends.pdf>

C.2 HSCIC's Social Services Activity³¹ and Expenditure³² Statistics

Background to the data

Community care is the process by which requests for social care help made to Councils with Adult Social Service Responsibilities (CASSRs) are translated, via assessment and care planning into appropriate services. Health and Social Care Information Centre (HSCIC) publishes a suite of statistical reports on adult social care in England. This includes the National Statistics, *Social Services Activity (Activity)*, which presents the number of referrals made to CASSRs, and the number of people receiving assessments, reviews and services funded by CASSRs. HSCIC also produces the National Statistics, *Personal Social Services Expenditure and Unit Costs (Expenditure)* which provides information about the money spent on adult social care by the social services departments. The underpinning information from councils is used in their day to day management of service users care plans and subsequent care packages.

HSCIC operates under the authority of the *Health and Social Care Act 2012*. This broadened the role of HSCIC, and established it as an Executive Non-Departmental Public Body. Following a social care data review, HSCIC is changing its social care data collections – to be implemented in 2014-15. These include a new Short and Long Term (SALT) Return.

HSCIC consults CASSRs and then confirms the arrangements for future collections in annual and quarterly letters to the Directors of ASSs. As part of the programme to replace the adult social collections, HSCIC circulates a monthly newsletter among councils to share ideas and examples of solutions through case studies from councils. It has seconded LA staff members to work on the implementation programme.

Care workers, social workers, care managers, and council administrative workers maintain operational databases used in the day to day management of service users' care plans. A care record is created for each service user and carers when they are assessed for social services. This is a record of the individual's needs and any services they require access to, this can also be referred to as a care plan. These records/care plans are maintained and updated when care packages for users are reviewed and changed.

³¹ <http://www.hscic.gov.uk/catalogue/PUB13148>

³²

<http://www.hscic.gov.uk/searchcatalogue?productid=13760&topics=1%2fSocial+care%2fSocial+care+expenditure&sort=Most+recent&size=10&page=1#top>

Quality assurance

Councils routinely carry out data cleaning exercises on their systems to ensure the data are current (for example removing deceased clients). Councils use the data to monitor the quality of service delivery provided by their care teams.

Two secure data transfer systems are used in the adult social care returns – Data Depot and Omnibus. Omnibus has built-in validation routines to check the validity and completeness of the submitted data. A validation report for each CASSR highlighting any potential issues and the reports is emailed to the appropriate contacts. The expenditure information received via Data Depot also undergoes some system consistency checking, such as, identifying blank cells and auto-sums.

HSCIC's quality assurance activities include: system checks of missing data; investigations of missing information through contacting individual councils; examining internal consistency within and between tables; consistency checks over time and examining the plausibility of the data. For example, HSCIC compares the Omnibus data for the number of new service users to the number of service users receiving services and to the rate per population to check the plausibility of the data.

The results of HSCIC's checks are presented in an annex in *Activities* – it is a detailed presentation of the completeness of the main variables for each council and highlights any particular limitations with the data. It also explains if estimation was required to make up for missing information and the approaches taken.

Audit

HSCIC has not specifically documented the nature of audits conducted within LAs but it told us that it thought it might be possible to find out about such arrangements through its annual supplier questionnaire. Its Statement of Administrative Sources for these data highlights the routine quality assurance within councils but does not refer to their audit arrangements.

For the past two years HSCIC has prepared a review of the quality of the nationally submitted health and social care data. This comprehensive report³³ presents some examples of good practice by suppliers, as well as the steps taken within HSCIC to quality assure the data. It identifies some key messages about the importance of good quality data across the health and social care sectors and emphasises the difference between data validity and accuracy. It recognises the need to extend 'auditing of data quality beyond Payment by Results to other areas where poor data quality could impact most on direct clinical care'.

³³ <http://www.hscic.gov.uk/article/2021/Website-Search?productid=12280&q=quality+of+nationally+submitted+health+and+social+care+data&sort=Relevance&size=10&page=1&area=both#top>

There are a number of external sources of information about councils' social care data that can provide useful evidence to support judgments by statistical producers and users of the suitability of the data and factors affecting their quality, for example, such reports may highlight whether any concerns have been raised about the completeness or accuracy of the underlying information. The investigations may also indicate the overall effectiveness of the safeguards established by the local authorities in managing their services, and, in turn, provide some assurance or raise a concern about the quality of the administrative data. These sources include:

- information collated and published by the councils themselves about their systems and services through internal audits and local accounts of social service delivery
 - Health Quality Improvement Partnership (HQIP) provides support for providers to conduct social care audits³⁴ and has released the Quality Accounts Resource³⁵ to provide information for care providers planning their quality accounts submission
- information collated across councils:
 - by the Local Government Association (LGA) across councils about the social care services through its LGA Inform website³⁶ and
 - the Adult Social Care Outcomes Framework (ASCOF) indicators³⁷ collated by HSCIC on behalf of Department for Health
- the results of social care audits by and on behalf of regulatory bodies such as the Care Quality Commission (CQC)
 - CQC now regulates and inspects care providers against minimum standards of quality and safety³⁸
- national reviews e.g. by National Audit Office which published an overview report³⁹ on adult social care in England in March 2014

The findings from these reviews will not necessarily lead to quantitative estimates of quality but can provide a richer body of evidence to inform judgments about the suitability of the administrative data for use in producing and using official statistics.

Findings

- HSCIC has established detailed quality assurance processes to validate the data received from councils
- It provides clear and detailed guidance provided to councils in supplying the adult social data

³⁴ <http://www.hqip.org.uk/social-care-audit-frequently-asked-questions/>

³⁵ <http://www.hqip.org.uk/social-care-guidance-and-resources/>

³⁶ <http://www.local.gov.uk/about-lginform>

³⁷ <http://ascof.hscic.gov.uk/>

³⁸ <http://www.cqc.org.uk/content/fresh-start-regulation-and-inspection-adult-social-care>

³⁹ <http://www.nao.org.uk/report/adult-social-care-england-overview/>

- HSCIC’s social care statistical team has some knowledge of the data issues associated with the data from their contacts with the suppliers. An explanation of specific issues identified during quality assurance is included in the quality sections of the *Activity* bulletin
- HSCIC’s understanding of local issues was enhanced through the secondment of local government staff
- However HSCIC does not receive information about the steps taken by councils to audit their data and has not reviewed reports from NAO, local accounts or care quality accounts
- The statistical team told us that it could seek information from the councils about their audit arrangements using the questionnaire sent to all suppliers each year
- HSCIC also has quarterly and annual letters setting out the changes planned to collections and can use these to flag issues around audit and assurance
- HSCIC’s annual quality report provides an important organisational statement about the issues that affect the operational uses of health and care information as well as the recording and of data underlying official statistics
- LGA Inform provides a useful information source for HSCIC’s statistical team to learn about steps being taken by councils to improve social care delivery such as through peer review and local accounts
- HSCIC could follow up with some of the main organisations to find out whether data issues were identified during self-assessment and peer reviews and the efforts made to improve local accounts

Wider lessons

- Identify audit and assurance issues at an Institutional level:
 - An organisational level review of the quality of data collections provides important insight into the factors that affect the accuracy of administrative data for both the statistical producer and data supplying bodies, as well as for users – these can be issues that are of central concern to operational (e.g. clinical) uses, as well as for statistical needs
- Develop co-operative arrangements with data supplier organisation(s):
 - Establishing close ties with data supplier bodies is particularly important in developing a good understanding of the issues affecting data quality
 - The secondment of staff into the statistical producer body – and potentially from the statistical producer into the operational setting – is a valuable means of improving knowledge about the data collection and operational issues that may affect the quality of the statistics
- Determine the credibility of the data supplier organisation(s) to provide accurate and reliable data:

- Identifying the wider governance arrangements, such as through benchmarking schemes and quality accounts, will provide both statistical producers and users with a better appreciation of the issues affecting the accuracy of the statistics

C.3 HSCIC's Hospital Episode Statistics⁴⁰

Background to the data

Hospital Episode Statistics⁴¹ (HES) is a data warehouse held by the Health and Social Care Information Centre (HSCIC) which contains details of all admissions, outpatient appointments and Accident and Emergency attendances at NHS hospitals in England. The data are collected during a patient's time at hospital and are submitted to allow hospitals to be paid for the care they deliver. The HES system is designed to enable secondary use – that is, for non-clinical purposes – of the administrative data. Provisional HES statistics are produced and published by HSCIC on a monthly basis. The final annual reports are published as National Statistics.

The HES statistics have a wide range of users within HSCIC, by policymakers in the Department of Health (DH) and across a range of public and private health organisations. HSCIC also provides secure access to the underlying data to around 200 trained users from a range of organisations, such as government departments, hospital trusts and public health observatories. These users have access to pseudonymised⁴² record level data.

Quality assurance

Healthcare providers record patient data in a range of local patient administration systems, to support the care of the patient. The data are submitted to the Secondary Uses Service⁴³ (SUS) data warehouse. The raw data are then made available to commissioners and also copied into a database for processing. At pre-arranged dates during the year, data are extracted from SUS and then sent to HSCIC for processing and loading into the HES warehouse.

The HES data quality team in HSCIC validates and cleans the extract and derives new items. The team discusses data quality issues with the information leads in hospital trusts who are responsible for submitting data to SUS. The roles and responsibilities within HSCIC are clear for the purposes of data quality assurance, i.e. to assess the quality of data received against published standards and report the results of those assessments, but there is no central sign-off mechanism for the data submitted to SUS.

HSCIC has a well-developed data quality assurance process for the HES data, once the extract is received from SUS. It has about 700 data suppliers and uses an xml schema to ensure some standardisation of the data received. The use of the schema means that the dataset has to meet certain validation rules before it can be

⁴⁰ <http://www.hscic.gov.uk/hesdata>

⁴¹ <http://www.hscic.gov.uk/hes>

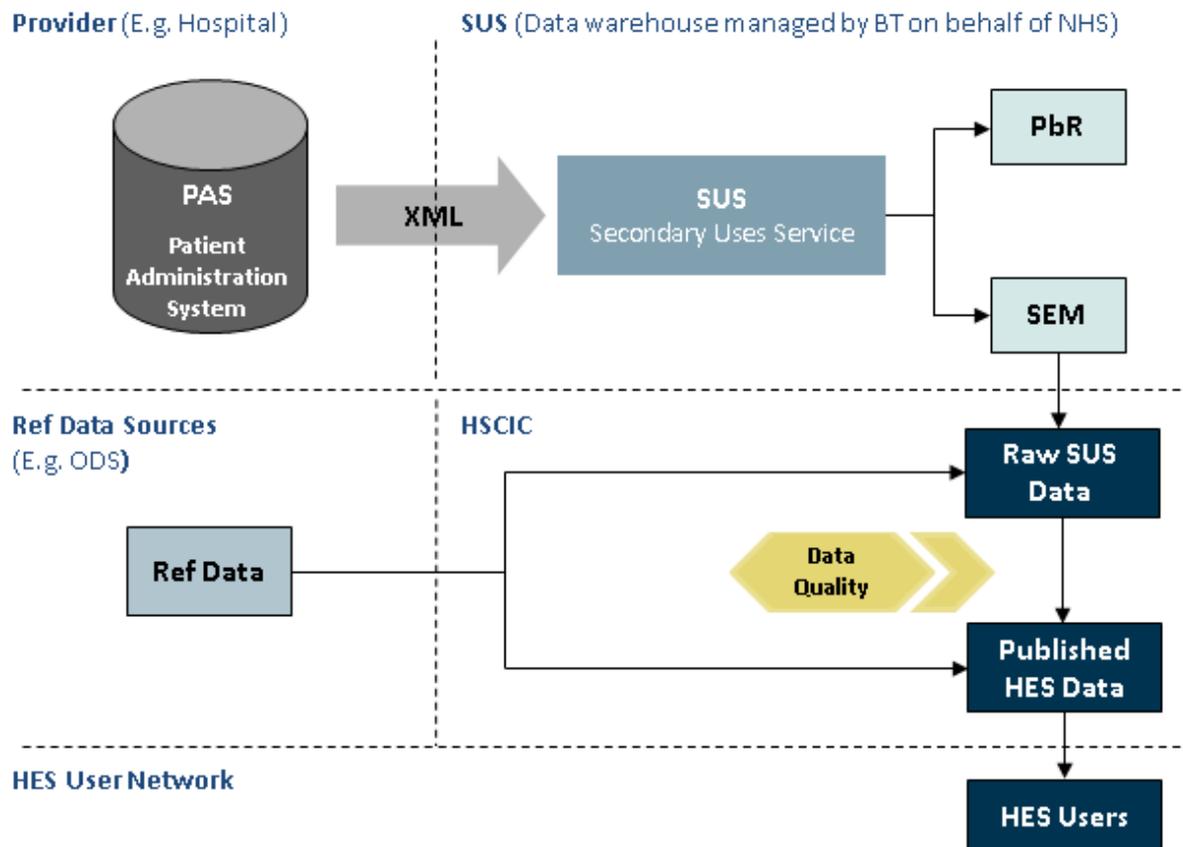
⁴² <http://www.hscic.gov.uk/dles>

⁴³ <http://www.hscic.gov.uk/sus>

submitted to SUS. HSCIC leads on the schema changes and consults the data suppliers about proposed changes.

HSCIC publishes detailed information about how the analysts collect and process the data used in the HES reports. This includes useful process charts to show the flow of data through the system (see Figure C3.1 below).

Figure C3.1: Illustration of the HES processing cycle



PbR = Payment by results extract

SEM= SUS extract mart

ODS = organisation data service (a website⁴⁴ with code reference library for health and social care organisations)

The range of guidance⁴⁵ also includes:

- Information about the data quality checks and data cleaning that the analysts carry out on the data after it is extracted from SUS, in order to prepare them for publication. HSCIC publishes the cleaning and derivation rules⁴⁶ for HES that have been developed over time

⁴⁴ <http://systems.hscic.gov.uk/data/ods>

⁴⁵ <http://www.hscic.gov.uk/media/1366/The-HES-processing-cycle-and-HES-data-quality/pdf/>

⁴⁶ <http://www.hscic.gov.uk/article/1825/The-processing-cycle-and-HES-data-quality>

- A further document⁴⁷ outlines the data quality checks performed on SUS and HES data when the Commissioning Data Set (CDS) is submitted via xml to SUS and must pass the validation to be accepted, the SUS Business rules and the checks carried out on the **Payment by Results (PbR)** dataset
- Data quality information for each year to date HES dataset is published in the monthly HES data quality report⁴⁸. These are published at the same time as the provisional year to date HES data. The statisticians can only check the validity and format of the data and not whether it is accurate, as accuracy checking requires a level of audit capacity and capability which the HSCIC does not currently possess
- HSCIC's second annual report *The Quality of Nationally Submitted Health and Social Care Data*⁴⁹ highlights issues around the recording of the underlying data that are used for HES, as well as examples of good and poor practice

Audit

HSCIC told us that its role is clearly defined in the *Health and Social Care Act 2012*⁵⁰ and that this doesn't extend to a regulatory role over the health care providers. For that HSCIC relies on the organisations that have such powers, such as, CQC or Monitor⁵¹.

HSCIC told us that the number of data suppliers (about 700) means that it is not possible to audit each individually. HSCIC said that the responsibility for the quality of the data submitted to SUS lies with the suppliers, even when held in the HES data warehouse.

DH currently contracts an independent external auditor to audit the accuracy of the data submitted to SUS, and used to calculate payment for activity, against that recorded in patients' notes (see section below). The HSCIC does not currently have the level of audit capacity and capability to do the same for data that is not used to calculate payment for activity, even though it may impact on the quality of care. However, HSCIC identified the importance of addressing coding issues in its annual quality report, *The Quality of Nationally Submitted Health and Social Care Data*. It noted that clinicians often see little direct value in their roles in using HES data and in ensuring good quality data, so it is often left to coders to interpret patient notes when coding. The quality report highlighted the concerns about quality of health records previously stated by Robert Francis QC in his report on the Mid Staffordshire

⁴⁷ http://www.hscic.gov.uk/media/13655/Data-quality-checks-performed-on-SUS-and-HES-data/pdf/Data_quality_checks_performed_on_SUS_and_HES_data.pdf

⁴⁸

<http://www.hscic.gov.uk/searchcatalogue?productid=14625&q=title%3a%22Provisional+Monthly+Hospital+Episode+Statistics%22&sort=Relevance&size=10&page=1#top>

⁴⁹ <http://www.hscic.gov.uk/catalogue/PUB11530/second-annu-data-qual-rep-2013.pdf>

⁵⁰ <http://www.legislation.gov.uk/ukpga/2012/7/contents/enacted>

⁵¹ <http://www.monitor.gov.uk/about-your-local-nhs-foundation-trust/regulatory-action/action-were-taking-nhs-foundation-trusts>

NHS Foundation Trust Public Inquiry⁵² and in Dame Fiona Caldicott's review of information governance⁵³.

From 2007 to 2012, the Audit Commission delivered a data assurance programme for PbR, the NHS tariff for paying acute hospitals. The work included an audit programme⁵⁴. The audits reviewed the key data that underpin payment in the NHS's PbR system. From 2013/14 the PbR data assurance framework is managed by the DH on behalf of NHS England and Monitor. DH commissioned Capita to deliver the PbR data assurance framework⁵⁵. In these audits Capita examined patient notes and compared these to what was recorded and submitted to SUS. While there are gaps in terms of what is included, for example, Clinical Commissioning Groups are to be given the option about what aspect of their data can be examined, it contains some information that could be usefully interrogated by the statistical producer teams as they consider the credibility of the data for statistical purposes.

Findings

- HSCIC publishes clear guidance for users about its data quality assurance processes on the patient level data in Secondary Uses Service (SUS), used to compile the HES statistics. This includes useful process maps
- HSCIC told us that it has clear internal processes for this data quality assurance work
- HSCIC has good knowledge of the quality of the data submitted to SUS for all fields where data standards exist for values and formats. However, aside from the information available via the PbR audits, HSCIC does not have knowledge of the accuracy of the data
- HSCIC is aware of wider audit work that could impact on some of the data presented in HES, but to date it has not used this as part of the quality assurance process
- HSCIC's annual data quality report provides examples of good and poor practice among health and social care data suppliers
- Overall HSCIC's quality assurance of data submitted to SUS for HES is thorough; however, its statistical teams could consider what other sources of information can be used to support their judgments about the suitability and accuracy of the data for producing official statistics

Wider lessons

- Gather information about audit arrangements from external data suppliers

⁵² <http://www.midstaffspublicinquiry.com/>

⁵³ <https://www.gov.uk/government/publications/the-information-governance-review>

⁵⁴ [http://www.chks.co.uk/Payment-by-Results-\(PbR\)-Assurance](http://www.chks.co.uk/Payment-by-Results-(PbR)-Assurance)

⁵⁵ <http://www.chks.co.uk/userfiles/files/PbR%20Key%20Findings%20Report%202013.pdf>

- The statisticians should consider other sources of existing audit that could provide them with some reassurance about the quality of the underlying data that they are using to compile the National Statistics
- Collate organisational information relating to the quality of underlying data in a central depository
 - HSCIC could create a central knowledge store to hold information about clinical audits across different health topics. This would allow the statisticians working on different topics easy access to these audits and less duplication of effort for each team to search for them. The statisticians could then consider, for each set of statistics, if any of the stored audits could be appropriate for their topic area or provide any insight into the underlying data.

C.4 ISD's Prescription Statistics in Scotland⁵⁶

Background to the data

The Information Services Division (ISD) a division of National Services Scotland, part of NHS Scotland, publishes a range of prescribing statistics. They cover different aspects of prescriptions for example dispenser remuneration, the number of items dispensed and costs of drugs.

Dispensing contractors, i.e. community pharmacists, dispensing doctors and appliance suppliers, are contracted by NHS Scotland to provide a service to the population of Scotland. To ensure drugs are available for dispensing when a patient arrives with a prescription, dispensing contractors buy prescription drugs in advance and then seek reimbursement for the drugs they dispense. NHS Scotland publishes statistics about its payments to dispensing contractors, these are remuneration for the service they provide and reimbursement for the products they dispense. These statistics are based on the data generated when a prescription is created, dispensed, and for which a claim is made.

As prescription expenditure covers in excess of £1 billion per annum and around 15% of the total NHS general revenue allocation in Scotland, there is strong interest from the Scottish Government and NHS service providers for information to ensure the clinical and cost effectiveness of these treatments. Prescribing data are also used for policy development, target monitoring, and for medical research such as clinical trials and epidemiology.

*Scottish Drug Tariff*⁵⁷ provides information about the prescribing, dispensing and reimbursement of medicines and appliances on primary care prescriptions. Types and value of dispensing fees are agreed with the Scottish Government and set annually. Details can be found in *Scottish Drug Tariff* and in Primary Care circulars⁵⁸ issued by the Government. Payments are derived from information gathered by Practitioner Services Division (PSD) in NHS Scotland, after the pricing of prescriptions has taken place. Reimbursement payments are made monthly to pharmacy contractors and will vary according to activity and claims. Some remuneration payments are set annually; some are now updated quarterly and some are set on a varying scale. Payments to dispensing contractors are made by PSD on behalf of the NHS Boards (the regional bodies responsible for the delivery of healthcare in Scotland). The data are collated and managed through a national payment system. The data warehouse brings together prescription, dispensed and patient information. Most prescription records (95%) generated by a GP in a GP practice have the patient's Community Health Index (CHI) number.

⁵⁶ <http://www.isdscotland.org/Health-Topics/Prescribing-and-Medicines/Community-Dispensing/Prescription-Cost-Analysis/>

⁵⁷ <http://www.isdscotland.org/Health-Topics/Prescribing-and-Medicines/Scottish-Drug-Tariff>

⁵⁸ [http://www.sehd.scot.nhs.uk/pca/PCA2013\(P\)21.pdf](http://www.sehd.scot.nhs.uk/pca/PCA2013(P)21.pdf)

The vast majority (about 92%) of prescriptions dispensed in the community are written by GPs, of which 99% are supported by an electronic prescription message. The remainder are written by other authorised prescribers such as pharmacists, nurses and dentists. When a GP writes a prescription, two forms are created; a physical form that the patient can take to a pharmacy and an electronic form (called an eMessage) that creates the first record of the payment cycle. When the patient then takes the physical prescription to a pharmacy, the pharmacist can access the eMessage and check the paper record against the eMessage for prescribing purposes. The dispensing contractors submit paper forms to PSD twice a month and PSD has ongoing access to the eMessages. The dispenser should also submit a dispensed e-message claim as well as supplying the paper copy of the prescription form for scanning. The prescription scanning process takes place on a monthly cycle.

PSD supplies information to the Prescribing Information System (PIS), which holds information on all NHS Scotland prescriptions dispensed within the community and claimed for payment by a pharmacy contractor (i.e. pharmacy, dispensing doctor or appliance supplier). The data include CHI numbers, prescriber and dispenser details, costs and drug information where available. Some research has estimated that these latter prescriptions account for around 6% of all prescriptions issued to patients. It is not possible to determine from payment data how much of the medicine dispensed to patients is actually taken in accordance with dosage instructions or why the medicine was prescribed.

The statisticians in ISD receive a dataset from PSD on a monthly basis. The data are stored in the Prescribing Information system (PIS), and ISD publishes quarterly official statistics about the payments made to dispensing contractors. ISD also run PRISMS, a web-based application which gives limited NHS Board access to prescribing information for prescriptions dispensed in the community from April 2004 onwards. The information is held centrally and the system is updated monthly. PRISMS can be interrogated to provide reports by individual prescriber, practice, locality, Community Health Partnership, NHS Board and for Scotland as whole. There are around 500 active NHS Board users of PRISMS and around 100 NHS Board users of PIS.

Quality assurance

PSD told us that there are approximately 5.3 million prescription claims per month covering 8.5 million items, supported by electronic messages for 4.3 million prescription messages for 6.5 million items respectively. The data go through several stages of checking on the payment system before being submitted to PIS for the statisticians to compile the statistics. These include checks of the pharmacy paper forms against statutory regulation, in-built validation checks of eMessages against the Dictionary of Medicines and Devices rules, and checks of what was dispensed compared with what was prescribed as recorded in the eMessage. There are also IT

validation steps in place for the loading of data into PIS; and checks of output by ISD when the data are first loaded (with tolerance levels for expected output compared to earlier time periods). In addition to this, the high use of the data increases the identification and amendment of anomalies in the data.

PSD has established business rules that determine item re-imburement. If the electronic claim meets these business rules, then the claim can be automated. About 63% of item re-imburements on the system are now automated, covering about 70% of GP prescribing. If the electronic claim fails to meet the business rules then the claim is sent for manual checks and item re-imburement. Pharmacy contractors now have access to the data through online reporting and can also check their payments.

Routine monthly checks are carried out by PSD on a random sample of approximately 5% of prescription payments (the flow of prescription data and related checks are summarised in Figure C4.1 below). These check all data captured for payment and the accuracy of the payment calculation and have a target accuracy of 98% which is routinely met.

Audit

In addition to the routine financial audit described above, PSD told us that there is an annual, independent audit of the service. Also Community Pharmacy Scotland⁵⁹, the professional body representing community pharmacists, carries out further checks on the data. This is ongoing, and is generally on a random sample.

Audit Scotland has published some work in this area in 2003⁶⁰ and in 2013⁶¹. The latter report looked mainly at the financial implications of dispensing different types or makes of drugs, rather than auditing the payment process. While not specifically highlighting quality issues, it provides useful contextual information to support the interpretation of the statistics.

The statisticians at ISD told us that they do not generally explain the audit arrangements in the prescription statistical reports. They monitor quality issues by developing reports that can identify anomalies. They also share the data with NHS Boards who provide a further level of scrutiny.

Findings

- The prescribing data goes through a range of checks and balances but little information is published about this detailed assurance by PSD
- The statisticians also do not publish information about the financial checks that could be used to support the statistics

⁵⁹ <http://www.communitypharmacyscotland.org.uk/>

⁶⁰ http://www.audit-scotland.gov.uk/docs/health/2003/nr_030626_supporting_prescribing_km.pdf

⁶¹ <http://www.audit-scotland.gov.uk/media/article.php?id=226>

- The statisticians could also look to include information about the independent financial audit that is carried out annually on these data

Wider lessons

- Fully document audit arrangements and identify the implications for the statistics:
 - Data that are used for payment purposes are likely to have detailed checks and quality assurance carried out on them, in addition to regular independent financial audit
 - The statisticians should look at these sources of audit and consider if they provide evidence of the robustness of the data and associated implications for the production and use of the statistics
- Gather information about audit arrangements from external data suppliers
 - The statisticians should obtain sight of these audit reports to satisfy themselves of:
 - a) the relevance of the audit for the statistics and/or
 - b) the quality of the underlying data
- Develop a clear process map:
 - Where the underlying data are subject to complex route through the process cycle, the statisticians should consider the inclusion of a process map to identify the stages in the data cycle and the data checks that occur at each stage

- 4 Access to ePAY system for; NSS staff - manual queries and system checks; pharmacists - checks of own payments and dispensing
- 5 Validation rules: based on form types; dispenser type; item rules; business rules & tolerances
- 6 eMessages reconciled against scanned forms
- 7 eVADIS reference data: drug and prescriber reference files. Continually checked and maintained
- 8 CHI reference information: patient details, demographics and GP information. Continually checked and maintained
- 9 Warehouse brings together prescription, dispensed, and patient information; ~ 8.5 million items per month. 95% have CHI number (i.e. linked to a specific patient). Suite of reports run after each data load (monthly) for load checks; tolerances; exception reporting
- 10 Routine quality assurance and checks by analysts/statisticians of information and statistical products/publications produced from PRISMS and PIS
- 11 Adjustment payments made to correct over/under payments (Dec 13: 430 items out of ~ 9 million total number; value of £45k out of £8m total)

In addition to the checks and processes above, random sampling and external audit to assure NHS Board payments and ensure appropriate controls and checks are in place.

C.5 DWP's Work Programme Statistics⁶²

Background to the data

DWP's Work Programme is the government scheme in Great Britain to assist people who are long-term unemployed into sustained employment. The service is provided by employment support organisations through 40 contracts with 18 prime providers. These providers work with a larger number of sub-contractor organisations. Claimants are randomly assigned by the local Jobcentre Plus to a prime provider in their area.

The providers are paid when the claimants complete defined periods in work (usually after 6 months, but 3 months for those that are hard to place e.g. ex-prisoners) receiving the job outcome payment. The providers then can receive further monthly payments for each additional month in employment – i.e. sustainment payments.

DWP uses the statistics to monitor the performance of the programme. It has a business plan transparency indicator that it reports to HM Treasury every quarter. This metric is included in the official statistics bulletin. DWP monitors the performance of the providers through minimum performance levels which are also presented in the statistics bulletin.

Other main users of the statistics are: Parliament (through select committees), local government, employment support organisations (including the providers themselves), the third sector and the media. They use the statistics to hold the government to account, as well as to benchmark local delivery.

The data on referrals to the Work Programme are taken from Jobcentre Plus's administrative system (Labour Market System – LMS) used for administering customer claims and includes the claimants characteristics and claim details. The data on attachments to the Work Programme, as well as information about the payment of job outcome and sustainment payments, are submitted to DWP by the prime providers only (i.e. the sub-contractors do not supply data to DWP) for payment purposes. These records are submitted through the Provider Referral and payment System (PRaP) (see Figure C5.1).

Quality assurance

DWP statisticians undertook a large number of quality assurance checks during the early development of the Work Programme statistics to test: the reliability, completeness and level of disclosure of individual variables; the levels of duplicate, missing or contradictory information; and the consistency across computer systems and with management information. They also carried out a number of other investigations to determine the suitability of the data sources for use in producing

⁶² <https://www.gov.uk/government/collections/work-programme-statistics--2>

official statistics on the Work Programme. These included analysing the trends and variation in characteristic, time series and geographical breakdowns, as well as examining trends and differences in post-payment adjustment factors.

This quality assurance found no issues; it showed the data were robust, consistent with management information, and suitable for publishing. Individual variables were complete and consistent with existing sources, comparisons across systems were always within a 0.1% tolerance, including across key breakdowns and time series.

From this validation, the DWP statisticians then developed routine quality assurance procedures, including: data cleansing rules; checks on the scale of revisions as payment data are updated on the database, and automated checks against tolerances for a sample of tabulation tool tables. DWP reported⁶³ that individual variables remain complete and consistent with existing sources, comparisons across systems remain predominantly within a 0.1% tolerance, including across key breakdowns and time series.

Audit

a) Financial audit:

All Job Outcome payment claims are subject to an 'off-benefit check' before payment. This involves an automated check to match participant information on DWP's Customer Information System, to ensure that participants for whom Job Outcome payments are claimed are not claiming benefit.

The automated off-benefit check has a window of 15 days in which the check is applied, to allow for minor discrepancies between the details of the provider's claim and the details on Departmental systems. Job Outcome payment claims that fail this automated check are removed from the system (unless they can be validated) and not paid. Claims which pass the off-benefit check are released for payment, and are then subject to further post payment in-work checks.

Post-payment validation is performed every month to strengthen the controls against fraud and error in the Job Outcome payments reported to DWP by the Work Programme providers. This process involves selecting a sample of 33 claims per contract (for the latest 3 months statistics to be published) from the total population of Job Outcome payments that passed the automated off-benefit check and that were subsequently paid in that month. The sample is drawn at random from the population of Job Outcome payments paid in the sampling period and is of large enough size (across 3 months) to enable DWP to extrapolate error rates. DWP reviews each extrapolation round to ensure that it operated as intended.

⁶³ <https://www.gov.uk/government/publications/work-programme-official-statistics-background-information-note>

The sample is matched against HMRC P45 data to validate employment. Those that fail the HMRC check are validated by confirming employment with either the employer or the individual. Job Outcome payments that are found to be invalid are used to calculate the error rate which is extrapolated from the total population. The results of 3 rounds of validation (one for each month) are brought together every financial quarter to provide quarterly error rates used in the official statistics.

The primary purpose of the error rate is to extrapolate financial recoveries against all payments made to a contract in the extrapolation period based on the error rate, rather than for the sampled claims alone. Once the percentage of error has been calculated from the sample, the error rate is applied to the total paid to providers for the relevant three-month period, and the provider is then required to pay this back to the Department.

The error rates for the 40 contracts are used to derive adjustment factors which are then used to rate the official statistics to reflect final Job Outcome payments made to providers. The adjustment factor is derived using the number of the Job Outcomes which fail the post-payment validation process divided by the total number of Job Outcomes sampled. This ratio is applied to Job Outcomes (less the sample and those already validated) to adjust the official statistics.

Once the validation process has been completed, Work Programme providers have the opportunity to challenge its results. Time is allowed for providers to challenge and for DWP to assess and arbitrate any challenge. This process can take up to approximately 3 months, so that the official statistics on Job Outcomes for some providers may be revised slightly in the following quarterly release. DWP says that the affect of these revisions have so far been minimal.

The end to end post-payment validation process takes approximately 5 and 1/2 months to complete. The routine sampling, checks and production of error rates take just over 1 month and these are performed on the previous 3 months Job Outcomes payments.

b) External Audit

NAO conducted a data assurance audit⁶⁴ of DWP business plan transparency indicators and the data systems underpinning these metrics in 2012-13. NAO identified that DWP had undertaken a good assessment of the risks at the various stages of the collection process, including assessment of likelihood, impact and mitigating controls. It found that there are comprehensive desk instructions for each stage for the quality assurance and data validation processes; automated pre-payment checks on claims entered by providers including the 'off-benefit check' by DWP; and independent quality checks by a different team in DWP.

⁶⁴ <http://www.nao.org.uk/report/2012-13-review-of-the-data-systems-for-the-department-for-work-and-pensions/>

This audit concluded that:

- DWP has set out its strategy for the management and use of information, including its policies and standards to protect data and customer privacy, as well as how it re-uses information more effectively
- It has a data quality policy but needs to improve clarity over roles and responsibilities
- The Work Programme transparency indicator's system was rated as **'adequate but some improvements could be made'** – NAO highlighted that the latest version of the PRaP system had not been internally audited after its implementation.

DWP told NAO that it planned to do an audit of the PRaP system in the next financial year – it told us that it has completed its testing.

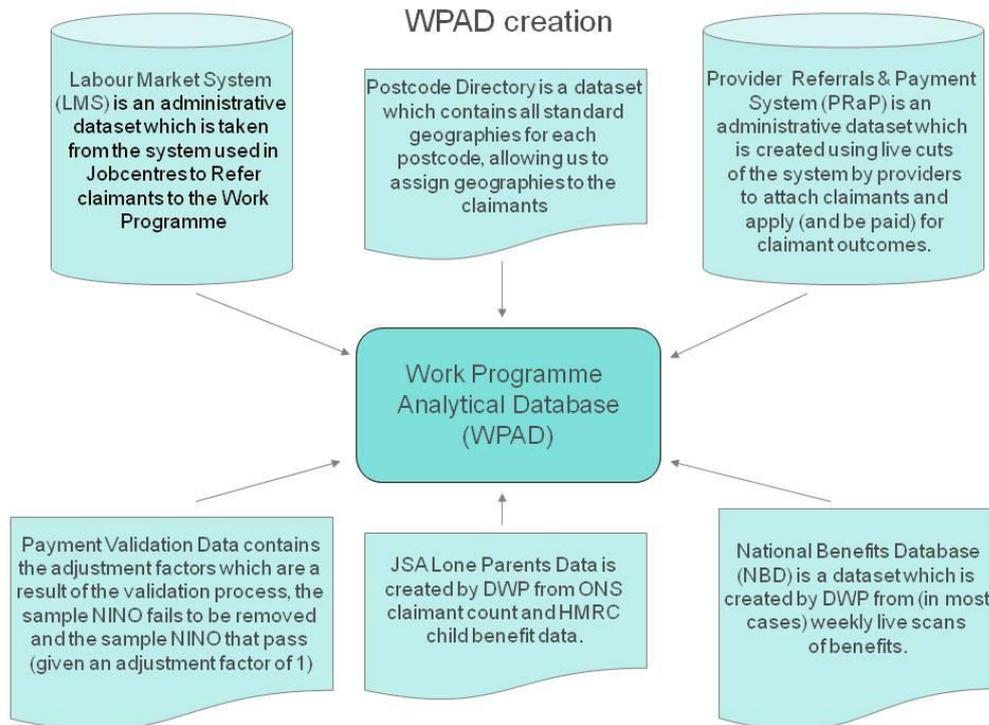
Findings

- The statistics are largely based on financial data received by the department for payment purposes
- Detailed validation and quality assurance both manually and through automated system checks
- Manual checks involve direct contact with claimants to verify benefit status
- NAO conducted audit of the data systems and rated them as 'adequate but some improvements could be made' – recommending improvement in the clarity of the Data Quality Policy regarding roles and responsibilities and that DWP conducts an audit of the provider data system (PRaP)

Wider lessons

- Fully document audit arrangements and identify the implications for the statistics:
 - DWP has published information on its financial validation and quality assurance
 - It sets out the method to adjust the underlying data to prevent bias from delays to payment information undergoing validation and the impact on the statistics of operational delays in validation of payments
- Develop a process map:
 - DWP could provide further clarification of these arrangements such as by using the flow chart for the creation of the Work Programme Analytical Dataset, to indicate the main quality assurance steps and relevant quality indicators, e.g. PRaP error rate, % completeness of records, linkage rates
- Undertake internal audit of systems and processes for administrative data:
 - DWP provides reassurance of the robustness of the provider data by conducting its internal audit of the PRaP system

Figure C5.1: Diagram showing the various data sources feeding into the Work Programme Analytical Database



C.6 Health and Safety Executive's Injury Statistics⁶⁵

Background to the data

Administrative data on specified fatal and non-fatal injuries, occupational diseases and dangerous occurrences are collected under RIDDOR (*Reporting of Injuries, Diseases and Dangerous Occurrences Regulations 1995*). The regulation places a legal duty on employers to report certain workplace incidents to the relevant enforcing authority: HSE; local authorities; or, the Office of Rail Regulation (ORR).

HSE took over responsibility for the RIDDOR notification system in September 2011. Reporting by employers to HSE became predominantly online, using newly designed forms and online guidance. To produce the RIDDOR injury statistics, data are extracted each quarter from the RIDDOR database by the HSE statisticians, and a copy made for statistical purposes. HSE stores injuries information on an operational database for its area of enforcement. The database is maintained continuously. HSE combines data from its database with data provided by LAs and ORR to produce the RIDDOR injury statistics. These are National Statistics. The fatal injury figures are released quarterly, while the non-fatal injury figures are published annually.

Quality assurance

HSE makes a series of validation checks on the statistical data, looking into implausible data values, and making any necessary adjustments prior to publication.

In the case of fatal injuries, additional checks are made by correlating with additional sources of information, usually related to an investigation of the incident. All fatal injuries are investigated by HSE operational staff (the HSE inspectorate). A number of different sources of information are used in these investigations, just one of which is a formal RIDDOR notification from employers.

There are far more reports of non-fatal injuries than fatal injuries. A relatively small percentage of the major injuries are investigated further by the HSE Inspectorate; otherwise, there is no further detailed investigation of the events. The information provided by the employer when reporting an injury is taken largely at face value. HSE says that it has no cost-effective way to verify the information provided; however, it has commissioned a statistical audit of a sample of non-fatal injury records (see audit section below for further information) which provides some information on the scale of under-reporting.

The main quality assurance strategy involves a number of aspects of checking and validation; for example, conducting: system validation checks in the RIDDOR database to prevent incorrect data entry; data cleaning checks on data items in the statistical dataset; and consistency checks on the raw data, such as, comparisons

⁶⁵ <http://www.hse.gov.uk/statistics/causinj/index.htm>

against previous datasets. HSE also asks the relevant LA and ORR to confirm the accuracy of their injury data.

Audit

The regular review of injury records by the HSE inspectorate provides one type of audit of the injury data. Two other types of audit have been conducted by or on behalf of HSE: internal audit; and the statistical audit of non-injury data. HSE also compares the reporting of non-fatal injuries obtained through RIDDOR with the self-reported survey data from the Labour Force Survey (LFS). It commissions these survey questions to gain a view of work-related illness and workplace injury based on individual's perceptions and also presents these statistics in its annual statistical outputs.

a) Internal audit

In March 2012, HSE's Internal Audit team reviewed the RIDDOR system. This review was initiated by the statistics team following the transfer of responsibility of the injury notification collection system to the team in September 2011. The Internal Audit team's review focused in particular on how the process of reporting of fatal and major injuries was working and examined: wrongly allocated reports; backlogs of unallocated reports; the clarity of guidance documents in relation to reporting of incidents; the clarity of information provided in some aspects of the reports; and, the experience of local offices and HSE switchboard in responding to enquiries.

The team identified some areas requiring improvement. Following the audit, the statistical team:

- introduced some improvements to the online reporting form;
- provided guidance to assist the completion of the form;
- changed the review process to determine whether an incident was reportable or not; and,
- worked with the front line staff to understand better the potential impact of malicious reports on the injury statistics.

Internal Audit also conducted a follow up review in April 2013 to assess progress in addressing the required improvements and determined that appropriate actions had been taken.

b) Statistical audit of non-injury data

HSE commissioned a survey to check the information recorded on RIDDOR by speaking with the injured employees about the event. It enables HSE to better understand the issues that impact reporting on non-fatal injuries and the potential biases that occur as a result, as well as to provide information on the amount of time taken off work for reporting to Eurostat.

Based on the information obtained from interviewing around 2,000 injured people from a random sample of records of non-fatal injuries reported by employers, the survey found that:

- For injuries reported as major by employers:
 - 90% were confirmed as major
 - 10% were found not to reach the threshold for a major injury (that is, were over-reported)
- For injuries reported as over-7 day injuries by employers:
 - 60% were confirmed as over-7 day
 - 23% were under-reported and subsequently found to be major injuries
 - 17% were over-reported (i.e. were below the threshold required)

Overall, however, the survey concluded that the method was sufficiently rigorous to produce robust estimates of the average number of working days lost to workplace injury per worker to meet Eurostat's needs.

Findings

- Audit arrangements are integrated within RIDDOR process for fatal injuries through investigations conducted by the H&S inspectorate
- All notifications are reviewed by the H&S inspectorate, including non-fatal injuries
- Two reviews of the RIDDOR system were conducted by Internal Audit – the findings have been implemented by HSE
- A statistical sample audit has been conducted for HSE of non-fatal injuries – the results have not yet been published but has identified levels of under- and over-reporting
- HSE supplements data about non-fatal injuries by using self-reported data from the Labour Force Survey
- Some aspects of the audits have been published by HSE, and its quality assurance steps are explained alongside the statistics, but fuller information about the range of audit and investigation carried out on the injury data could provide further insight about the quality of the statistics to users
- HSE has developed a process map of its data supply arrangements
- HSE has not asked about audit arrangements in the LAs and ORR

Wider lessons

- Develop a process map relating quality assurance and audit arrangements into the operational and data supply arrangements:
 - HSE developed a process map to better help it understand any areas of weakness in the RIDDOR system (see Figure C6.1)
- Fully document audit arrangements and relate these to the implications for the statistics:

- HSE has documented the strengths and weaknesses of the statistics but did not specifically refer to the audit arrangements within these and their implications. Audit is not mentioned in the 'Background Quality Report' for the injury statistics
- Publish a summary of the findings of the statistical sample audit and make clear the implications for the injury statistics in relation to use
- Gather information about audit arrangements from external data suppliers:
 - HSE obtained some information for the data it had collected but not for the data from other suppliers
- Use triangulation – identify alternative data and information sources e.g. from surveys – to verify accuracy of the data
 - HSE compares the administrative data on injuries with self-reported statistics from the Labour Force Survey

Figure C6.1: HSE process map for incidents reported by employers under RIDDOR (as at May 2014)

