Information Services Division (ISD)

Overview of Data Quality of Sources and Outputs

Document Control

Version	0.9
Date issued	13/01/2017
Author(s)	Jill Ireland
Other related documents	
Comments to	NSS.NSSstatsgov@nhs.net

Document History

This is a dynamic document which aims to include quality commentary on every dataset in ISD.

Version	Date	Comments	Author(s)
0.1	02/02/2016	Initial draft	Jill Ireland
0.2	06/03/2016	Changes to content following discussions with Richmond Davies, data management staff and analytical teams	Jill Ireland
0.3	21/04/2016	Review by Richmond Davies – changes to structure and other minor changes	Jill Ireland
0.4	13/05/2016	Additional content to include information on data suppliers' quality assurance	Jill Ireland
0.5	16/05/2016	Additions to content following discussions with Richmond Davies, data management staff and analytical teams	Jill Ireland
0.6	15/06/2016	Additions to content following feedback from UKSA on Public Health Assessment	Jill Ireland
0.7	12/01/2017	Additions to content following further feedback from UKSA on Waiting Times Assessment	Jill Ireland
0.8	13/01/2017	Additions to content following discussions with Richmond Davies	Jill Ireland
0.9	13/01/2017	Amendments to ensure document reads well	Jill Ireland

Contents

1. Background

Following extensive consultation, the United Kingdom Statistics Authority (UKSA) published a Standard for the Quality Assurance of Administrative Data for Official Statistics on 29 January 2015. It sets out the principles the UKSA will deploy as it forms its judgements about whether Official Statistics merit the National Statistics status, which indicates that the statistics are trustworthy and meet the highest standards of quality and value.

Quality assurance of administrative data is more than simply checking that the figures add up. It covers the entire statistical production process and involves monitoring data quality over time and reporting on variations in that quality. The Quality Assurance (QA) Matrix (shown in Appendix 1) helps the assessors and producers to determine the types of assurance and documentation required to inform users about the quality assurance arrangements for administrative data. It guides the judgement about the suitability of the data and identifies examples of practices that meet the appropriate level of assurance.

The QA matrix identifies the evidence required for achieving each of the four levels of assurance (A0-none, A1-basic, A2-enhanced and A3-comprehensive), across the following areas of practice:

- Operational context and administrative data collection
- Communication with data supply partners
- QA principals, standards and checks by data suppliers
- Statistical producer's investigation and documentation

2. Overview of the Approach to Quality in ISD

Principle 4 of the *Code of Practice for Official Statistics* details the best practice of ensuring that official statistics are produced to a level of quality that meets users' needs, and that users are informed about the quality of statistical outputs.

However, principle 7 of the Code states that statistical producers should seek to balance quality (e.g. timeliness and accuracy) against costs (i.e. data supplier and statistical producer resources), taking into account the expected uses of the statistics (e.g. official statistics *versus* management information for the needs of different types of users).

UKSA's publication *Quality, Methods and Harmonisation* argues that there is no universal definition of quality and that it varies in concept in relation to different procedures and products.

ISD adopts an approach to quality assurance which is captured in the quality feedback cycle (Figure 1):

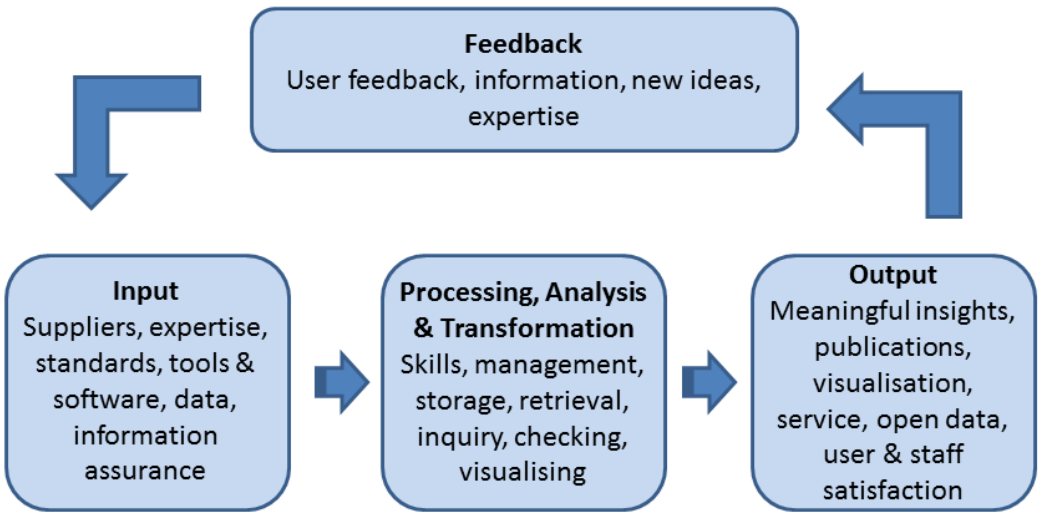

Figure 1: Quality Feedback Cycle

Figure 1 above shows that, in response to the needs of users (e.g. policy makers, healthcare providers, researchers, the public, etc.) or legal imperatives, data suppliers in health and social care provide data to ISD that comply with agreed standards which arrive through various routes and are captured and stored. Data governance is applied, which normally involves collaborative working between data suppliers and data management professionals, information analysts and IT and software engineering professionals. The work includes processes that ensure usability, integrity and security of the data as well as procedures to deal with data quality issues and their resolution.

Data management professionals in ISD monitor and advise on data quality aspects of data governance such as completeness, conformance with standards and definitions, and work closely with data suppliers to resolve and prevent data quality issues.

Analysts retrieve extracts of the data from within the corporate data warehouse or national databases using statistical and reporting tools and techniques. They also have quality assurance processes built into their activities which include techniques for testing the cleanliness of the data and adhering to the organisation's analytical checking guidelines. They add value by working efficiently to make sense of the data and apply descriptive, diagnostic and predictive analytical techniques to provide meaningful insights and intelligence that meet the needs of users and contribute to health and wellbeing of people in Scotland.

The outputs include the responses provided for our wide range of customers who use our freedom of information and research request service, our regular Official and National Statistics publications, contributions to Scottish parliamentary questions, data contributions to UK-wide and international compendia, and our rapid turnaround services for the media and other members of the public.

While transforming inputs into outputs, processes are in place to monitor accuracy, timeliness, accessibility, clarity and coherence which are all dimensions of quality. There are robust statistical disclosure control processes which ensure information governance principles are adhered to (i.e. data protection, accountability, integrity, compliance, retention,

availability and disposition). As shown in figure 1, a feedback mechanism exists whereby feedback is continually sought from users to improve the collection, analysis, presentation and dissemination of our statistics. This is a transparent process that ensures trust in statistics and secures maximum public value as required by protocol 1 of the Code of Practice for Official Statistics.

Data quality assurance underpins all of the work ISD undertakes on behalf of NHS Scotland, as the quality of national data and data outputs is a key requirement of all those who use our outputs to improve health and wellbeing in Scotland.

3. Quality Assurance of Data Inputs

The ISD Quality Assurance process in the lifecycle of each publication involves collaboration between analysts, data management professionals, data suppliers, information governance professionals, IT experts and key stakeholders with a continuous improvement mechanism using user feedback.

ISD aspire for all our statistical publications to meet level A3 - comprehensive assurance across all four areas in the Quality Assurance Matrix (Appendix 1). However, ISD recognises that due to (a) ongoing continuous quality improvement work at source, the pace of which cannot always be influenced by ISD, and (b) work in progress to digitally transform and rationalise ISD's outputs via its Smart Publishing Improvement Programme, many of our outputs will only meet level A1 or level A2 quality assurance.

As a result of the range of quality assurance methods in place at source (e.g. validity, consistency and accuracy checking, coding, adhering to guidance documents, senior NHS Board official sign-off, reporting quality concerns, site visits from ISD, embedded ISD staff within NHS Boards, ISD helpline and forums for discussing data capture issues, and practitioner and counter fraud monitoring of systems), ISD remains confident that data suppliers understand the need for robust quality assurance processes and have processes in place even though these processes generally differ across NHS Boards. These processes at the data supplier end are regularly tested, wherever possible, by ISD analysts and data management professionals.

Further information in relation to specific datasets follow.

3.1 Scottish Morbidity Record (SMR) and other National Datasets

ISD's Data Quality Assurance (DQA) team is responsible for evaluating and ensuring SMR and datasets are accurate, consistent and comparable across time and between sources. Not all NHS Boards use the same administrative system, but they are all configured to supply the data fields which are consistent for national reporting purposes. Documentation exists to provide guidance for data suppliers at NHS Boards with regards to the system, specification and required format of the data. It includes flow diagrams that illustrate the end-to-end process and timetables for ensuring timely data submission to ISD. The team has established good working relationships with their contacts in NHS Boards and they expect regular questions on aspects of quality of the data collected at source. A flow

diagram illustrates how the DQA team uses the quality assurance database system to assess and monitor the quality of data from the data suppliers.

The DQA team has a key role in assessing data at source, ensuring the most accurate data are submitted and arranging for ISD's clinical coders to provide adequate training to NHS Board staff as required. Once the DQA team has completed an assessment through a combination of site visits and secure remote access to electronic data and physical records, their findings are analysed and reports produced for each NHS Board which are sent to Board coding managers, chief operating officers and medical directors. These are then followed up by series of meetings at NHS Boards to discuss specific findings and actions for Boards.

All data supplied from NHS Boards are verified, and ISD's Data Support & Monitoring team liaise with data providers to ensure data is validated against national standards, and to enable any issues to be identified at an early stage. Where appropriate, any issues identified are then returned to the contact at the NHS Board for correction and re-submission. Accuracy and completeness are monitored and discussed regularly via email and also during site visits with the data suppliers at NHS Boards. There are also regular discussions between data management and the analysts, in terms of the regular SMR Analyst Forum (SAF) and also topic specific meetings pertaining to individual data marts. The team provides a helpdesk support to data suppliers, monitors completeness and timeliness of submissions, investigates data quality issues, carries our data validation checks, and works with suppliers to anticipate and address issues. They provide a service for the following national datasets:

- A&E data from emergency departments and minor injuries units - monthly aggregate and patient level data
- A&E waiting times data from emergency departments - weekly aggregate level data
- RTT – Referral to Treatment waiting times
- SMR00 - Outpatient data
- SMR01 - Acute inpatient and day case data (including geriatric long stay data SMR01_1E)
- SMR02 - Maternity inpatient and day case data
- SMR04 - Mental Health inpatient and day case data
- ISD(s)1 - Summary statistics of resources and activity in hospitals and other care settings
- Child Health data for preschool, school and immunisations
- Breast Screening data
- System Watch - weekly data from NHS Boards, Health Protection Scotland (HPS), Scottish Ambulance Service (SAS), NHS 24 and National Records of Scotland (NRS)
- Smaller datasets on a targeted basis

Changes to the data or definitions are required to go through a change management process and are discussed during the quarterly Change Control Meeting attended by representatives from data monitoring and analyst teams. If approved at this stage, the next step is to discuss at the External Change Control meeting which consists of representatives from NHS Boards, in addition to internal representatives.

National audits of the data are carried out, during which all NHS Boards are visited to assess the accuracy of data submitted to ISD. During 2015/16 and 2016/17, the Data Quality Assurance (DQA) team at ISD carried out a quality assurance of SMR04 data submitted to ISD with the aims of:

- Determining the accuracy and completeness of recording of selected administrative and clinical data items in line with national standards
- Determining if national clinical coding standards are being appropriately applied, highlighting and addressing any areas of confusion, identifying training requirements and sharing good practice
- Highlighting gaps in information being supplied to staff who record data items.

The exercise is nearing completion. The report will be published on the ISD website in the coming months. In addition, a quality assurance exercise of SMR02 data is currently underway.

National Data Standards are essential in order for the health and healthcare data held by ISD Scotland to be of high quality. The Data Advice Team is responsible for monitoring and managing national definitions, reference files and standards. The team ensures that the data are collected throughout Scotland according to the same classifications and rules and the data is interchanged between systems consistently, robustly and securely.

The data quality assurance, data support and monitoring and data advice teams are all part of the data management function in ISD and they work collaboratively with IT professionals, analysts and other internal and external stakeholders to ensure robust data governance for our national datasets.

With regards the SMR data capture at source, in all NHS Boards, patient information is collected as part of a hospital stay on a patient management system (PMS), or a patient administration system (PAS). The information recorded will relate to the stay and the patient's condition and may include information that is not sent to ISD. This information recorded forms the SMR dataset. An extraction process pulls the data, as per the current SMR rules, into an SMR input record which is then validated and submitted to ISD. Extracted records have a unique identifier attached to each record.

The SMR record is a bi-product from information that is already recorded by hospital staff with additional coding added to indicate the patient's diagnosis and any procedures performed. When the record is submitted to ISD there are additional derived items added. ISD's clinical coding service provides support across all NHS Boards to ensure consistency of coding and adherence to the coding standards.

Not all NHS Boards use the same administrative system, but they are all configured to supply the same data fields to ISD, to ensure there are no differences across areas in the recording of the data, and reduce the potential for any sources of bias and error. Comprehensive testing of the SMR extract functionality is required before live data submissions can take place. ISD and the NHS National Services Scotland's IT Customer Support Desk supply further advice and clarification for any queries in relation to the testing and submission process.

Ahead of submitting data to ISD, colleagues who provide data at the NHS Board level, need to undertake appropriate training. Standard operating procedures, PowerPoint slides, guidance documents and definitions manuals are in place to provide guidance on the entry of data to populate PMS or PAS. ISD has provided comprehensive documentation to the data suppliers to aid the submission of SMR records, and detailing the required format. In addition, training is also provided for anyone inputting ICD10 codes (International Statistical Classification of Diseases and Related Health Problems 10th Revision - list generated by the World Health Organisation (WHO)). Each person tasked with entering the data will have their own unique login details, to ensure there is an audit trail. Staff access to the electronic system will be limited to the areas that are pertinent to them.

The following example highlights how NHS Boards tend to react quickly to identify DQ errors at source. An NHS Board identified some quality issues affecting their SMR records since implementing a new data capture system. They quickly developed a suite of Data Quality reports to understand the impact of errors on SMRs. The internal Data Quality Project Group discovered impact was minimal. They also concluded they needed to increase the number of employed Data Quality Support Officers working within the NHS Board. The NHS Board then undertook a recruitment process to fill the new vacancies. Throughout the process, the NHS Board kept ISD informed. They also developed a Standard Operating Procedure (SOP) on how to code mental health admissions. During several meetings, ISD was satisfied the NHS Board had reacted in a timely manner to ensure future SMR data submissions retained integrity.

Each NHS Board is sent a timetable to show the last date for submission of SMRs for each month. This is based on a six week target. On the next working day ISD takes a snapshot of SMR data on the national file to enable the timeliness and completeness of the data to be measured.

To minimise any risks to data quality, all data supplied is verified. The source data is validated locally and goes through additional re-validation once submitted to ISD. Following the additional validation, a report is generated outlining any issues, which are returned to NHS Boards to rectify. This is usually done on a daily basis. The process is that the NHS Boards update the records queried and these are picked up when the data is next extracted. Additionally, checks take place to look for overlapping stays and duplicates in the SMR data.

With regards to the accuracy targets - all assessments use the ISD 90% standard as the target for accuracy on any data item. This is noted on the ISD website:
http://www.isdscotland.org/Products-and-Services/Data-Quality/How-We-Work/
Further information regarding the calculation of accuracy rates, and the basis for the 90% target, can be found on the following ISD webpage:
http://www.isdscotland.org/Products-and-Services/Data-Quality/Methodology/accuracy-rates-for-clinical-coding.asp

The NHS Boards supply a coding report from their local system each month, which includes the number of records submitted to ISD. This is checked against the SMR numbers recorded at ISD. The completeness is then calculated, based on the differences between

those numbers. This records any backlog in the data. The completeness figures are discussed during meetings with the NHS Boards.

The completeness data and timeliness charts (including details of backlogs) are also published on the ISD website:
http://www.isdscotland.org/Products-and-Services/Data-Support-and-Monitoring/

ISD analysts and data management colleagues take the opportunity to undertake site visits to gain a greater understanding of the process of capturing data at source and inputting into the SMR systems.

As noted previously, the Data Quality Assurance (DQA) team is within the Data Management function in ISD. The team conducts national audits of SMR data and other datasets. The audit process involves remote access to electronic data at source and site visits to NHS Boards to assess the accuracy of SMR care episodes that have been submitted to ISD by checking them against physical case files and IT systems at source. The teams not only assess the accuracy of clinical coding but also the suitability of local information handling procedures and identify any training requirements for NHS Board information staff.

There is clear agreement with data providers about what data will be provided, the frequency, through which medium, and by whom. The clinical coders on the ground and the regular site visits have provided a good appreciation of the context in which the data are collected including patient management or administration system performance, and the data providers accept that the quality standards agreed with ISD meet the needs of the statistical outputs that ISD produces. ISD also has the safeguards (e.g. communication channels, data advice and data quality assurance processes) in place to monitor quality and there is therefore generally a low to medium risk of data quality concerns. ISD is of the opinion the SMR datasets meet level A2 enhanced assurance. This is because the resources are not available to carry out very frequent formal data quality assurance exercises for all the SMR data sets. However, specifically for SMR04, ISD is of the opinion this should be assigned A1 basic assurance. This is because, although SMR04 is currently undergoing a QA audit, the previous formal QA audit was over 15 years ago. After the new audit report is published the level of assurance will be reviewed.

Generally, the public and media interest profile is medium as evidenced by requests for information or clarification and media headlines. There is also, in the main, medium political sensitivity and low to medium levels of risk of quality concerns.

3.2 National Smoking Cessation Database

Partnership Action on Tobacco and Health (PATH) was a joint initiative between ASH Scotland, NHS Health Scotland and the Scottish Government (SG). The aim was for the key partners to develop and roll out best practice across key areas of training, data collection and evaluation as follows;

- SG tasked PATH to lead on the development of data collection for smoking cessation services.

- Following consultation, a 'minimum dataset' (MDS) & detailed guidance were produced.
- NHS Boards were required to collect MDS, from clients attending services, for national monitoring.
- Smoking cessation services incorporated MDS into local data collection.

In terms of the development of the Smoking Cessation database, ISD was tasked to develop and manage a web-based IT system, or 'database' to aid the capture and analysis of client data. It was a requirement to enable secure electronic transfer of MDS information to ISD for national analysis, plus additional data items and functionality to meet local needs.

Data are submitted to the database from specialist smoking cessation services via two methods:

- Direct key entry by registered system users from each health board via secure encrypted connection. The system captures real time data.
- File upload from the national Pharmacy Care Record (PCR) system. This system is utilised by community pharmacies across Scotland through a secure encrypted connection. Records are extracted from PCR and uploaded on a weekly basis to the smoking cessation database.

In terms of direct key entry, the dataset comprises both MDS and non MDS fields. The MDS fields are centred on the quit attempt and each stage of follow-up, which are at 1, 3 & 12 month post quit date. Validation rules, created by the ISD Data Management Team, are linked to these fields and, in some instances, cross field validation exists. If any of the core MDS fields are incomplete, a warning is issued on saving the record, which highlights all the missing fields. Guidance on using the system and completing the MDS fields are available as PDF downloads from the system.

The PCR is a web-based application to support pharmacies when providing Chronic Medication Service (CMS) and selected pharmacy-led services. Smoking cessation is just one of the modules available for their use. The model has inbuilt validation akin to the smoking cessation database, devised after extensive consultation between ISD and ATOS, the PCR developer. Practitioner and Counter Fraud Services of NHS National Services Scotland have electronic systems in place to monitor the quality of data captured and submitted from pharmacies in order to prevent fraud. Files are received on a weekly basis from the PCR system. The files will contain data pertaining to client quit attempts and/or follow-ups at both 1 & 3 month post quit date. PCR does not provide 12 month post quit follow-up data. Pharmacies will not have direct access to the smoking cessation database. Once the files arrive at PHI, they are stored in a holding folder until scheduled time for upload.

In terms of the data at source, the evidence ISD has received from NHS Boards indicates the following:

Pharmacy:

All pharmacists and their staff are required to undertake training. These include training on data requirements and the PCR system. In addition, the NHS Board runs face-to-face training on an ad hoc basis.

The Public Health Pharmacists also visits individual Community Pharmacies (CPs) to help and support CPs with regards to the Community Pharmacy Stop Smoking Service (CPSSS). There is also mentoring from the NHS Board Pharmacy Champions on the CPSSS when they make face to face visits on the 10-14 week follow-ups, or during visits to the CPs that each Pharmacy Champion covers. Throughout 2015 and 2016, CPSSS has been a priority area for all Pharmacy Champions.

There are various processes in place within the NHS Board to verify accuracy and to quality assure data from the CPSSS. Public Health Pharmacists chair local quarterly Smoking Cessation and Prescribing group meetings to discuss performance, prescribing levels, training requirements etc.

Support on specific areas of the CPSSS is also given. For example, the number of blank Pharmacy Care Records (PCR) records in CPs was investigated by one NHS Board. The NHS Board undertook a pilot with a sample of CPs in order to reduce the numbers of such blank PCRs.

As all patient data is submitted via the electronic PCR as part of the National service specification, the NHS Board no longer sees any paper recording forms. Payment is not made unless the electronic form is completed with the minimum dataset requirements and submitted within the required timeframe. If the pharmacy does not submit their electronic form, they do not get paid – this incentivises accurate reporting.

Quit rates are fed back to pharmacies at least annually, usually more frequently. Within the NHS Board, the performance review group also receives information on numbers and quit rates and runs a watching brief. The payment verification team from NSS look at smoking cessation payments and numbers and highlight any abnormalities. A quarterly report is sent to the NHS Board from them.

Non-pharmacy:

Within the Smoking Advice Service, the data is firstly collected directly from the client by the advisor/referrer. This is then checked by administrative staff, before it is entered. Any queries regarding accuracy will be checked with the advisor/referrer/client. The weekly PCR upload errors received from NSS are also checked, and a response provided where required.

Quit rates, referrals and demographic information are looked at frequently, as part of the reporting for the Chief Executive's Letter (CEL), Health Promoting Health Service (HPHS) and Local Delivery Plan (LDP). Regular performance and review meetings are held within each NHS Board. Target performance data is provided regularly by data analysts within the NHS Boards to support these discussions. In addition, feedback is provided to pharmacies and those who refer also.

All stop smoking staff have received appropriate training to ensure data input accuracy, with ISD providing local information on missing or inaccurate data. All staff who handle client data have completed information governance / security training and those assessing patient information systems have signed the relevant disclaimers.

Data accuracy is reported via the Impact Performance Report for discussion at appropriate groups to establish processes and procedures to ensure data accuracy.

Once data is uploaded, the ISD Data Management team run additional validation checks, before uploading to the smoking cessation database. Any record that does not pass validation will be recorded as an error, which will be forwarded to the NHS Practitioner & Counter Fraud Services (P&CFS) helpdesk. The helpdesk will forward errors to the appropriate NHS Board, who in turn will contact the individual pharmacy contractor to resolve the issue. All NHS Boards are issued with guidance pertaining to handling errors.

ISD Data Management staff check completeness on a quarterly basis and contact NHS Boards regarding completeness and data quality issues. In addition, the staff will notify NHS Boards of any outstanding issues at the end of each calendar year. National smoking cessation coordinators meetings are held on a quarterly basis and provide updates pertaining to:

- Data quality/completeness
- System issues
- System changes

National smoking cessation coordinators meetings include representatives from Scottish Government, NHS Boards, NHS24, British Heart Foundation, and ISD analyst and data management teams. ISD analyst and data management team representatives also attend the Smoking Cessation Database Project Meetings, along with representatives from academia, the Scottish Government, ASH Scotland, NHS Health Scotland and territorial NHS Boards. Data management teams provide updates pertaining to system issues and system changes.

Responsibility for management of prisoner healthcare was initially with a non-NHS organisation who, at the time, provided performance management reports each month to NHS procurement and the Scottish Prison Service. From October 2011, the responsibility was transferred to NHS Boards. Collection of prison quit attempts and outcomes have been recorded on the ISD Smoking Cessation database since 2013.

Resourcing for the provision of NHS Smoking Cessation services and the collection of data regarding their uptake within prisons varies from prison to prison. In addition, care must be taken when comparing prisons, since there are different types of prison, each catering for specific types of offence prisoners, together with the lengths of sentence handled. Completion rates for the collection of data is improving, allowing the publication of the number of attempts analysed by prison, together with the number of quits and corresponding success rate at 1 month and 3 months displayed at Scotland level.

With regards the quality of data capture at source, the validation algorithms in the pharmacy care record system match those within the national smoking cessation database. These algorithms include cross-field validations. For example, pregnant women over the age of 44 are flagged for further examination in case of data input error. Pharmacy records account for 75%-80% of all quit attempts recorded in Scotland while NHS Board specialist smoking cessation services account for the capture of 20-25% of all quit attempts. These data are captured by pharmacists and pharmacy technicians who are professionals registered with the General Pharmaceutical Council and they have personal accountability to their professional bodies to ensure that they comply with high standards of ethics, conduct and performance, which include being honest and trustworthy. This is particularly important because pharmacies see a much larger number of patients and therefore, their support for each patient cannot be as intensive as specialist cessation services in NHS Boards who see fewer numbers of patients and have higher quit rate outcomes. ISD does not have the resources, and therefore has not sought permission, for their data management professionals to check the accuracy of live data capture during pharmacy-patient interactions. However, ISD is assured that data capture from the patients, including pregnant women who smoke, is of sufficiently good quality because of:

- Personal healthcare professional accountabilities
- Pharmacy premises and NHS Board training arrangements
- In-built system validation rules and warning flags
- Guidance on the national minimum dataset for smoking cessation built within systems which promotes consistency and has resolved late data submission issues
- ISD's support for reporting, discussing and resolving data quality issues with healthcare professionals (e.g. site visits and telephone and email consultations)
- ISD's quality assurance processes once the data arrive in ISD
- The continuous monitoring of systems by Practitioner and Counter Fraud Services for the purposes of detecting fraud in the NHS
- ISD's representation at the national smoking cessation network meetings

The level of risk of the data quality concerns is medium because high risk factors have been moderated by the safeguards as listed above. The public interest profile of the statistics is classed as low because the interest mostly lies within a niche user base (i.e. the Scottish Government and pharmacies) and media interest in Scotland is limited. The Chief Executive of ASH Scotland finds the statistics to have high value and has provided the following feedback in terms of how the Smoking Cessation publication is used within her organisation:

> "*ASH Scotland relies upon the integrity of ISD smoking cessation information, and their regularly revised statistics are immediately updated throughout our website and on our factsheets. ISD statistics underpin ASH Scotland strategies, work plans and funding applications as well as playing a key role in maintaining a solid evidence base on which to counter tobacco industry misinformation. We regard ISD statistics as the gold standard when communicating information to policy makers and to the media, and having this solid and trusted foundation enables us to work towards the Scottish Government's ambitious target of reducing smoking prevalence to less than 5% by 2034. In an economic climate which places severe pressure on the funding of interventions to address health behaviours such as smoking it is more important than ever to have good metrics and ISD presents them in clear, accessible formats upon*

which we have come to rely. In short, it is hard to envisage that ASH Scotland could function fully without the services of the Information Services Division (ISD) of NHS National Services Scotland" (June 2016).

Given that the smoking cessation statistics have medium quality concerns and low public interest, ISD is of the opinion that they met meet level A2 enhanced assurance.

3.3 Waiting Times datasets

ISD analyst teams attend regular meetings with representatives of NHS Boards and the Scottish Government to discuss data collection and the quality of the statistics. The waiting times datasets are particularly high profile because the analytical outputs from the data are underpinned by national waiting times targets and standards. The ISD website describes how waiting times are used by the major users in Scotland.

Detailed descriptions of the various waiting times datasets can be found on the waiting times pages in the ISD website. Each waiting times dataset section provides information about the data quality assurance, data source, data submission guidance, data definitions, associated national target or standard, and reporting. The waiting times analytical teams liaise frequently with data suppliers in the territorial NHS Boards on matters pertaining to accuracy, collection methods, data completeness and validation rule adherence. More specifically, the sub-sections below provide more information about the input and output quality of specific waiting times datasets.

3.3.1 18 Weeks Referral To Treatment (RTT) Waiting Times

The responsibility for delivering the 18 Weeks Referral To Treatment (RTT) target lies with the NHS Board that receives the initial referral, as this NHS Board will be responsible for agreeing with the patient and relevant clinicians the most appropriate pathway of care. In some cases patients may be initially referred to one NHS Board and then have an onward referral to another NHS Board for treatment.

Due to the constraints in current hospital information systems in linking all stages of a patient's journey to measure their waiting time as mentioned above, these statistics are presented on NHS Board of Treatment, the NHS Board where the patient's treatment was started. NHS Boards are working with ISD and Scottish Government to update systems in order to further improve whole pathway information capture to support the measuring and reporting against the 18 Weeks RTT target.

In relation to the other Waiting Times targets which relate to the 18 Weeks RTT pathway, the New Outpatient and Inpatient & Day Case waiting times information is in the National Waiting Times Data Warehouse at patient level. The Diagnostic Waiting Times information, in addition to the 18 Weeks RTT submission process, is provided to ISD by the NHS Boards as monthly aggregate submissions.

Aggregate data are submitted to ISD from all NHS Boards in Scotland using a standard Excel template. This Excel template was created by ISD in consultation with NHS Boards and the Scottish Government and is the template that all NHS Boards use to submit their

monthly 18 Weeks RTT information to ISD. The template has been refined over time in consultation with NHS Boards and ISD work closely with NHS Boards if they have any issues completing the template. In addition, ISD provide a guidance document to all data suppliers at NHS Boards on the data collection process to ensure consistency.

Since the data received by ISD are aggregated, patient-level information cannot be systematically validated by ISD. NHS Boards extract data from their systems and carry out quality assurance checks ahead of sending to ISD. In 2015, ISD Waiting Times conducted a consultation with all NHS Boards to review the 18 Weeks Methodology by asking the NHS Boards to provide details on a number of technical aspects, such as to provide details on the linkage method(s) used, how clock stops are identified, whether the validation method is automated or manual, etc. There is a mixture of both manual and automated validation undertaken within the NHS Boards to check the quality of the input source data. In addition, the data for each NHS Board are verified and signed off by the Chief Executive and the following quality questions are asked of the submitted data:

- Does the data, identify all patients on an 18 wks RTT pathway treated within your Board? If not, what is your estimate of the number of patients that you cannot report?
- Since the last publication, have you made any changes to your linking methodology?
- Are you aware of any new or ongoing data quality issues with regard to the 18 weeks figures that ISD should be aware of? Please give details.

The summary of the responses to these can be found in the data quality section of the publication to aid the reader.

The process for loading the NHS Board submission data into the ISD 18 Weeks RTT Database is owned by ISD's Data Management. The Data Management team follow the process as detailed in a Standard Operating Procedures (SOP) document to undertake data quality assurance checks on the information provided. The ISD Waiting Times team then additionally perform a limited suite of quality assurance checks on the data.

Given the nature of aggregated returns which ISD rely on for the 18 Week RTT data, there is a reliance on individual NHS Boards to ensure accuracy. ISD are committed to ensuring the accuracy of the data and to that effect, an ISD Audit of the 18 week RTT data commenced in November 2016, at the request of Service Access. In the five month period, between November 2016 and March 2017, a number of NHS Boards will be visited and 18 Weeks RTT information at source reviewed by ISD Auditors. ISD analysts and statisticians will also be involved in some of the NHS Board visits.

The main objectives of the audit are to obtain further detail on how the monthly ISD referral to treatment waiting times submission is calculated from patient pathway data at NHS Boards, observe the recording processes for waiting times data items and see how they are quality assured and linked. The ISD auditors will also ensure that the NHS Boards demonstrate that have documented operating procedures in place as well as adequate training for their relevant staff.

NHS Boards are working with ISD and Scottish Government to update systems in order to further improve whole pathway information capture to support the measuring and reporting against the 18 Weeks RTT target.

The 18 Weeks RTT statistics have medium quality concerns and high public interest. The level of risk of the data quality concerns is medium because there is clear agreement on which data will be provided, the frequency of provision, the means by which they are provided and the personal accountability within each NHS Board. ISD is satisfied that there is good appreciation of the context in which the data are collected given that RTT waiting times are so important for NHS Board capacity planning and decision making processes and also for meeting the national standard of delivery for at least 90% of patients. There is also recognition of the fact that the complete patient journey from referral to treatment can be difficult to measure. However, arrangements are in place, such as close scrutiny by both the Scottish Government and ISD on how the data are collected and methods of continually improving data collection and linking, given that the majority of NHS Boards have not been meeting the national RTT standard and that around 8% of patients journeys could not be fully measured. The data producers accept that ISD has active communication channels for monitoring aspects of data quality and technical issues affecting data extraction.

ISD is satisfied with the level of scrutiny relating to the 18 week RTT data, and the processes in place to ensure data are accurately entered and aggregated ahead of submitting to ISD. Given all the information outlined above, ISD is of the opinion that this dataset meets level A2 enhanced assurance.

Definitions and guidance for 18 Weeks RTT have been developed to help ensure that each patient's journey is measured fairly and consistently. The public interest profile of the statistics is high due to the high political sensitivity, media coverage of the statistics and the requirement for meeting a national standard.

3.3.2 Accident & Emergency (A&E) Waiting Times

Patient information on attendances at an Accident and Emergency Service is collected on a patient administration system (PAS). The information recorded will relate to the patient and details of their attendance and may include information that is not sent to ISD. An extraction process pulls the data into an A&E attendance record which is then validated and submitted to ISD soon after the end of each month. This information submitted forms the A&E dataset.
All sites that provide emergency care are required to submit data to the A&E datamart. There are two types of data submitted to the A&E datamart: episode and aggregate level data. All hospitals with Emergency Departments submit episode level data containing a detailed record for each patient attendance. Some smaller sites with minor injury units or community hospitals submit only aggregate level data as they do not have the information systems and support to enable collection of detailed patient based information.

NHS Boards have training and Standard Operating Procedures (SOPs) in place to help provide guidance on the entry of data to populate their A&E system. Each person tasked with entering the data will have their own unique login details, to ensure there is an audit trail. Staff access to the electronic system will be limited to the areas that are pertinent to them.

Data suppliers carry out a range of checks prior to submission of monthly A&E data, which include:

- checks of the internal consistency of the time fields recorded,
- checks of the validity of attendances recorded as being in breach of the 4 hour standards, and the associated reasons recorded
- check of consistency between related data fields e.g. referral source / arrival mode and discharge type / discharge destination

In terms of the data at source, feedback ISD have received from NHS Boards indicate that:
A number of NHS Boards perform weekly or more frequent checks of the data to ensure the submission of data to ISD is as accurate as possible. A report is available to front line administration staff that lists occurrences of the more common data issues. Additional reports are available daily for service managers to track breaches from the previous day, This was originally aimed at correcting discharge times where users had left default times in PAS, rather than changing to the actual discharge times. That is no longer an issue now, following training.

Whilst compiling a submission for ISD, a senior member of the analytical staff carries out a series of validations, again searching for occurrences of more common data issues, and sense checks against historic data values. The PAS system also provides an overall validation which will tell if records are incomplete prior to extraction.

After the data is accepted into the A&E datamart, a suite of performance and monitoring reports are produced. These are reviewed and any unusual or unexpected figures are raised with the NHS board for confirmation. If changes are required then they must be explained fully by the NHS Board and the data is resubmitted. A log is kept noting any changes that are necessary.

With regards the quality of data capture and recording at source, guidance documentation is provided to data suppliers with regards to system, specification and format. It also provides guidance on data entry and the validation applied when the data is submitted to ISD. In addition, it illustrates how the data suppliers can run error reports. The Data Support & Monitoring team within ISD provide helpdesk support to NHS Boards and other data providers who experience difficulty capturing, coding and submitting data. ISD Clinical Coding Tutors are available to provide support for NHS Boards and work with the NHS Boards to provide training where required and help with any changes which occur. In addition, each analytical team within ISD has their own generic mailbox for NHS Boards to submit any queries. This is monitored on a daily basis.

Once the data is captured on the NHS Board's patient administration system, data quality checks are undertaken. Some examples are the validation of correct codes carried out at data entry and checking of patient/person details against medical records.

Not all NHS Boards use the same administrative system, but they are all configured to supply the same data fields to ISD, to ensure there are no differences across areas in the recording of the data, and reduce the potential for any sources of bias and error.

Data are submitted to ISD using a csv file format through SWIFT-DVL one month at a time. If a new member of staff at an NHS Board wants to submit data, they require a SWIFT account. In order to obtain this, an email is sent to the Principal Info Analyst at ISD, outlining their name and job details. If the application is accepted, it is forwarded to NHS NSS Customer Services Desk (CSD) to action.

The data provider checks the content of the file submitted to ISD (e.g. correct number of records, format of file, header record). The submission file is tested by initial validation rules (e.g. correct file format, logic cross checks of data items, mandatory data items present and missing data items). Failure to meet any validation rule results in rejection of the submission and an automatic email error report detailing validation check failures is sent to the data provider. If there are no validation errors then data is loaded into datamart.

After the data is accepted into the A&E datamart, each month a suite of performance and monitoring reports are produced and made available to NHS Boards and the Scottish Government. The Data Support and Monitoring pages of the ISD website describe the full suite of data quality reports.

These are reviewed and any unusual or unexpected figures are raised with the NHS Board for confirmation. If changes are required then they must be explained fully by the NHS Board and the data is resubmitted. A log is kept noting any changes that are necessary. Any erroneous data can be resubmitted, and go through the initial validation and standard report reviews again. Once any queries are resolved, the data is considered fit for purpose.

There is no sign-off required for the data submission. By the 20th of each month, ISD send NHS boards a link to a secure webpage. This includes a management information report containing the data which will be published on the first Tuesday of the following month, and the suite of management reports described above. NHS Boards are asked in an email to highlight any concerns.

There is regular engagement with NHS Boards through site visits, emails and telephone calls. Data quality and completeness and submission of optional data items are discussed.

The A&E dataset review group, made up of representatives from ISD, NHS Boards and the Scottish Government began the process of reviewing the A&E dataset to ensure it remain fit for purpose for its stakeholders by collecting good quality, relevant data items without excess burden of data collection. By Spring 2015, the group had put together recommendations for a revised A&E dataset. A consultation exercise was delayed to await the results of the Royal College of Emergency Medicine (RCEM) / Health and Social Care Information Centre in England Emergency Care Dataset stakeholder consultation.

A&E statistics are published weekly and monthly. ISD has noted that all Boards invest resources to ensure that the data they collect and submit are accurate and valid because the monthly data are submitted to the A&E datamart which has in-built validation rules and data monitoring oversight. Whereas, the weekly A&E data are submitted as aggregated figures on a weekly basis via spreadsheets. These are two data supply routes, with the weekly one being aggregated information. The aggregated information in both agree with each other.

As there are around 1.6 million A&E attendances a year, some inevitable data errors have been identified and presented in the accompanying metadata document.(e.g. 7th June 2016 publication metadata) In this publication it is clear that the erroneous records are very few and they have no impact at all on the percentage compliance with the 4 hour standard either within a Board or at a Scotland level.

Given the number of safeguards such as the A&E user guide and recording manual for data suppliers, the validation rules that the data have to comply with, the two sets of complimentary A&E data recorded and submitted by different routes, the effective communication arrangements with NHS Boards, the scrutiny processes by ISD's data management team, the range of checks that ISD's data management team have verified that NHS Boards carry out before submission, and the common understanding of the need for accurate A&E statistics given well-publicised capacity pressures at A&E departments, the level of risk of data quality concerns is medium.

The public interest profile of the statistics is high, given that there is high political sensitivity as reflected by high media interest and reporting against a national standard. ISD is of the opinion this dataset meets level A2 enhanced assurance.

3.3.3 Child & Adolescent Mental Health Services (CAMHS) Waiting Times

When the Child & Adolescent Mental Health Services (CAMHS) data collection was first devised, the IT systems across NHS Boards were not set up to collect the data at patient level. Therefore, it was agreed to collect aggregate level data. The IT systems have developed since this work started and some of this information is now collected on NHS Board Patient Management Systems (PMS). However, there are still some services where the information is still collected by NHS Boards in Excel.

NHS Boards collate and submit aggregate level data to ISD in an Excel template. The template has evolved over time. The current template is set up to collect information on patients who waited during the month and information on patients waiting at the end of each month. This information (number of people) is collected in weekly time bands to allow calculation of the median and 90th percentile. A separate Excel sheet is set up for adjusted and unadjusted waits.

Given that ISD receives aggregated data from each NHS Board, this cannot be thoroughly validated by ISD. Derivations of the figures and data accuracy are matters for the individual NHS Boards. There is a great variation in who compiles the data in NHS Boards from administrative staff and information analysts to service managers. The NHS Boards do check the data to be submitted but again this varies from daily checks of the Waiting Times data to weekly or monthly checks. Checks prior to submission are carried out by a range of people: managers, clinical directors and heads of service. Some of the submitting NHS Boards have Standard Operating Procedures (SOPs) to assist them in the compilation of the data. Others are in the process of compiling theses. The NHS Boards discuss the data at team, management and performance meetings.

Data is quality assured within the NHS Boards at source. ISD has seen evidence to confirm that NHS Boards have a suite of reports to highlight common issues, such as missing/incorrect data, outliers, and list breaching patients. The reports are scrutinised for inaccuracies, and prompt the team's administrative support to source missing outcomes for the appointments from individual clinicians, and the team lead to review the waiting list and complete the section in the reporting template where NHS Boards are asked to detail action being taken to tackle long waits. Feedback regarding current status of breaching patients and the actions taken to ameliorate the situation are fed back to CAMHS Clinical Management Group. Updates are also provided to the NHS Board executive committee regarding progress towards the waiting times standard.

The senior analyst and service manager review all activity with the managers/management teams in each CAMHS locale and monthly analysis is undertaken ahead of submission, to identify any trends in the headline figures. Any anomalies are identified and investigated ahead of submission to the analysts at ISD.

ISD has programs set up to combine the NHS Board information into one file. Since this is aggregate level data, the analysis involves aggregating the numbers and calculating percentages waiting/waited and medians and percentiles. ISD also carries out quality assurance to sense check the data and liaise with NHS Boards to resolve any queries.

As part of the ISD publication process, NHS Boards are given five days to quality assure their data for publication. At this stage ISD will provide the NHS Boards with a list of data quality questions. These include questions related to changes to the data. This is to help ISD understand the data and also to check that any changes are not due to error.

The CAMHS waiting times data has been collected nationally since January 2010, although initially data were very incomplete and of poor quality. This led to the publication being labelled developmental. Since 2013 the completeness and data quality issues significantly improved and in February 2015 ISD decided the data was robust enough to remove the development label.

NHS Boards use different data systems to collect the CAMHS data. Therefore the internal checking processes for data at source are not identical across all NHS Boards. ISD have been sent an overview of the checking processes undertaken before the data is submitted, and as a result ISD are satisfied with the processes in place at the NHS Boards to ensure the accuracy of the source data. Prior to submission to ISD, the data also have to be checked by a senior officer within the NHS Board.

Training is in place for people entering the data at source. There is detailed guidance documentation outlining, for example, the specification for the aggregate data submission along with the definition for each data item and the timetable for data submission.

For each month and quarter when the data are submitted to ISD, ISD analysts carry out additional checks and will feedback any queries to NHS Boards. For aggregate returns, ISD is limited on the checks that can be carried out from ISD post submission. ISD liaise with NHS Boards to resolve any queries.

Additionally, as an additional quality assurance stage, NHS Boards are asked to check their data prior to publication. Access is given to a password protected area on the CAMHS waiting times website to enable the NHS Boards to view their data and benchmark against Scotland level figures. NHS Boards and ISD have shared responsibility to highlight any data issues prior to the final sign-off date.

ISD also routinely seeks clarification from NHS Boards where there may be large changes in numbers, unusual patterns in the data or changes in trends. These changes may be influenced by a variety of factors including service changes/reconfiguration or data recording changes.

ISD also works with NHS Boards to resolve local challenges that may be affecting data quality. For example, a small number of NHS Boards is not able to calculate CAMHS waiting times from referral to treatment. They are advised to use the second appointment as a proxy for treatment which is the guidance given by Scottish Government. Information on which NHS Boards are still developing their systems for this is detailed in the NHS Board level list of data quality issues. Some NHS Boards may not able to report on all service level tiers (i.e. individual CAMHS specialist, CAMHS multidisciplinary teams, and tertiary level services such as inpatient and day units). This may be because they do not provide services which fall under a particular tier or because they are still developing their systems to incorporate all tiers.

Sometimes, there are data completeness issues. Waiting times data are extracted from local administration systems which are updated frequently with information about appointments, attendances, etc. This may lead to different reported numbers of patients seen or waiting depending on the date the data were extracted. However, ISD has calculated that those differences normally equate to a relatively small proportion of total numbers of patients seen or patients waiting.

There Scottish Government-led CAMHS Implementation & Monitoring Group comprises representatives from the Scottish Government, ISD, NHS Education for Scotland and NHS Boards (clinical and information). This group is consulted if ISD needs to highlight any queries or inconsistencies on the use of the definitions. ISD would also highlight the latest data to this group and discuss any data quality issues.

With the safeguards and risk mitigation arrangements that are highlighted above, ISD is satisfied with the level of scrutiny arrangements relating to the CAMHS data including the processes in place to support NHS Boards to manage data quality issues at the data collection stages. The level of risk quality concerns is medium because of the safeguards and effective communication arrangements. ISD is also satisfied that data suppliers understand the context of the contribution of the data to its statistics publications particularly as spending in mental health is increasing in Scotland and there is an increasing interest in waiting times for psychological therapies. Annex 2 of the CAMHS publications provides detailed commentary on data quality and completeness for each of the NHS Boards. ISD is satisfied that Boards have a handle on data quality and completeness and have processes in place for addressing issues and highlighting any caveats that ISD should be aware of.

The CAHMS statistics have medium quality concerns and medium public interest as evidenced by the content of the metadata on data quality and the level of media, general public interest and political sensitivity. ISD is of the opinion that CAMHS statistics meet level A2 enhanced assurance.

3.3.4 Cancer Waiting Times

Cancer Waiting Times (CWT) data are collected by cancer trackers and administrative staff (including medical secretaries) across NHS Scotland which are then recorded onto the NHS Boards' cancer tracking systems.

ISD employs 27 cancer information officers who are embedded in NHS Boards all over Scotland and they are also involved in working with NHS Board staff to quality assure cancer data at source – cancer waiting times, cancer audit and cancer registrations. This arrangement provides an additional level of guarantee of robust quality of data at source.

Each NHS Board submits a monthly file to the Information Services Division (ISD) of NHS National Services Scotland, which contains episode-level records for each newly diagnosed primary cancer referral, which began treatment in the previous calendar month, for which they were the NHS Board of receipt of referral. Each record contains demographic information about the patient, key time points in the pathway (date of receipt of referral, date of decision to treat and date of first treatment), information on diagnosis and treatment, main reason for any adjustments to the waiting time, and main reason for any breaches of the 62 and 31 day standard.

The adjusted waiting time in days and the total of any waiting times adjustments in days is calculated by NHS Boards for each record. Each record also contains a flag which indicates whether it meets any of the exclusion criteria (complex clinical pathway, died before treatment or who refused treatment).

Guidance documents help provide training and support for staff who enter the data at source. Processes are in place if an NHS Board has definitional queries about cancer waiting times.

Each quarter, NHS Boards can resubmit monthly data to allow the most up-to-date information to be used for publication. This information is then validated and loaded onto the cancer waiting times database to allow data interrogation and reporting.

As part of the ISD publication process, NHS Boards are given a week to quality assure their data for publication. At this stage ISD will provide the NHS Boards with a list of data quality questions. These questions will ask about changes to the data. This is to help ISD understand the data and also to check that any changes are not due to error.

In 2013, a quality assurance audit was carried out on the cancer waiting times data. ISD analysts provided the quality assurance team with a sample of data and this data was checked against the case notes.

The report can be viewed, and it included the following recommendations:

- As many of the discrepancies were found in date fields, more care is required to ensure dates are recorded accurately.
- Care should be taken when submitting data for cross border patients.
- Improve communication between Boards and Networks in relation to cross border activity.
- Ensure most up-to-date guidance is utilised, with particular care taken when recording Date of Receipt of Referral for screened patients.
- Particular attention is required for cervical records which had the highest number of discrepancies.
- NHS Boards/Networks should be encouraged to undertake internal assessments regularly to improve the accuracy of recording, particularly for date fields. It is noted that that some NHS Boards have already implemented their own internal quality assurance processes

As part of the follow up, NHS Boards were asked to let ISD know if they have noted an improvement in the recording of date fields, and whether they had an internal quality assurance processes for the data in place. ISD was satisfied with the responses. Annual and, more often, monthly assurance checks are in place, to ensure that any discrepancies within the NHS Board are limited. Performance is monitored within trackers and discussed with the senior management team. To highlight some examples of work being undertaken, during 2016, the internal audit team within one of the NHS Boards conducted an exercise to evaluate the processes in place for ensuring the quality and robustness of data used to monitor and report on cancer waiting times. This is to ensure that any interventions can be made, where possible, to ensure the national targets are met by the NHS Board. The audit resulted in Category A – Good – There is an adequate and effective system of risk management, control and governance to address risks to the achievement of objectives.

Some NHS Boards have introduced a sign-off process to discuss the data prior to each monthly submission.

Training is provided for staff. For example, cancer tracker workshops are held every six months. During these workshops, cancer trackers can discuss any issues they have. Different scenarios are presented to test the participants.

Comprehensive guidance documentation is provided, which includes the specification (data items, format, values to use) of the data submission along with the definition for each data item. There is also information on who is responsible for submitting the data to ISD, how the data should be submitted, dataflow diagrams for the submission process, responsibility for accuracy and sign-off of the publication data, information on inclusion and exclusion criteria for the data and example scenarios to help NHS Boards with queries regarding the recording of cancer waiting times data.

This document is managed by the Cancer Waiting Times Data & Definitions Group. This group includes representatives from ISD, Scottish Government, NHS Boards and Clinicians. If ISD needs to discuss data quality issues with NHS Boards, this group would be used as a forum for this. It is, however, rare for ISD to have queries.

Validation of submitted data will be carried out by ISD on uploading of the data. This is a series of logic checks to see if the data is feasible, e.g. if the patient is born before they are referred for treatment. The SWIFT system will not accept the NHS Board submission if any of these logic checks fail. It should be noted, however, that this is not a measure of accuracy of the data.

A second level of validation occurs at ISD, and specific anomalies will be brought to the attention of each NHS Board. Examples of this are if the wait to treatment is over a year, or if the patient was referred a long time after they were treated.

The Data Quality Assurance (DQA) team on behalf of the Cancer Waiting Times (CWT) team carried out and completed a data quality audit for patients who fall under the 62 and 31-day standards in 2014. This was based on data from quarter 3 (July – September 2014). The audits help ensure that the data being recorded is accurate and that recording is consistent across Scotland. The DQA team supports NHS Boards to determine areas where data quality can be improved and helps them resolve any issues.

The DQA team has carried out investigations to determine:

- If correct codes are being used.
- If waiting times adjustments and exclusions (patient refused treatment, patient died before treatment or clinically complex) are being applied correctly and consistently, and whether there is any supporting evidence on the NHS Boards systems.
- If the rules and guidance are being appropriately applied.

More information on data quality including the results of the DQA can be found within the Data Quality section of the ISD Cancer Waiting Times web pages.

Given the information outlined, ISD is satisfied with the level of scrutiny relating to cancer data and the processes that Boards have in place to ensure data is accurately entered ahead of submitting to ISD. Further confidence on the source data was gained following the 2014 audit.

The level of risk of data quality concerns is low because over the years, data suppliers have become clear about the data they capture and supply to ISD including the specifications and format. There are also increasing levels of checks at source and communication between NHS Boards regarding cross-border patients. ISD's data management team provides support to the NHS Boards. Additional checks are run on the data received at ISD ahead of any analysis.

The public interest profile of the statistics is high and there is high political sensitivity as reflected by high media interest and reporting against a national standard. As a result, ISD is of the opinion that the Cancer data meets level A3, comprehensive assurance, in relation to the operational context and administrative data collection requirements, communication with data supply partners and QA principles, standards and checks applied by data suppliers, as outlined in the UKSA data quality assurance matrix.

3.3.5 Drugs & Alcohol (DA) Waiting Times

The National Drug Waiting Times Information Framework was introduced in October 2004. This included guidance and definitions on data items to be collected for drug waiting times. The aim of the framework was to give Alcohol and Drug Action Teams (ADATs) a consistent structure for local monitoring of treatment services. In February 2007, ISD provided ADATs with a revised data collection system that enabled ADATs to produce a wider range of reports and also provided the facility for ADATs to monitor data quality more easily.

Operational structures changed in October 2009, when the 22 ADATs were dissolved and replaced by 30 Alcohol and Drug Partnerships (ADPs). The Waiting Times Framework was designed to function at ADAT level. In October 2011 an improved Drug and Alcohol Treatment Waiting Times (DATWT) Database went live across Scotland. This was the first time that data on alcohol as well as drug treatments were recorded nationally, and the first time information was available on the full client journey from assessment to treatment. This database also facilitates reporting at ADP level.

The responsibility for provision of healthcare to prisoners in Scotland was transferred from the Scottish Prison Service to the NHS on 1st November 2011. ADPs are now responsible for the collection and submission of Drug and Alcohol Treatment Waiting Times in prisons.

Each service which has been set up must have a Service Level Agreement (SLA) completed. This agreement outlines the nature and purpose of Waiting Times data collection, in line with the aims of ISD. It constitutes a formal agreement between the data providers and ISD and details what is required by the data providers in terms of submitting reliable and accurate and timely data.

Data are submitted from source, either by direct key entry by registered system users or by file upload. ISD acknowledges that direct data entry increases error risk. However, ISD is satisfied that there is satisfactory compliance with the guidance documents that cover the data submission process, following several face to face discussions with NHS Boards. There are well established and agreed validation rules that ISD and the data suppliers understand. They show how the validation on the system relates to the dataset. There is regular communication between ISD and the ADPs to discuss quality of data capture. In addition, quality matters are discussed at the Drugs and Alcohol Data Action Reference Group. This group contains representatives from ISD, Scottish Government, NHS Health Scotland/National commissioned organisations (as appropriate) and Alcohol and Drug Partnerships (Technical Leads where possible).

A checklist has been produced to help ADPs ensure that the data in the system is fully updated and accurate by the date ISD make the extract. The data is signed off by the ADP Waiting Time Lead Co-ordinator on behalf of the ADP Chair & the NHS Lead prior to data extraction.

In terms of the data at source, feedback ISD have received from Alcohol and Drug Partnerships (ADPs) indicate that local systems used in some areas were developed with strict validation to ensure that data being entered met with the DATWT dataset. A number of regular checks are performed to ensure the accuracy of the waiting times data as far as possible. A number of the commissioned services have manual data entry procedures, so

the NHS Boards acknowledged a high degree of scrutiny is required to ensure data is maintained accurately.

Once all information is uploaded to the DATWT database, the ADP run several reports to review the waits that have been entered, focussing on on-going long waits in particular to ensure the national standard is met. Reviews are undertaken for waits that have breached the 21 day target, to check that these are not errors or poorly recorded data. Individual records are reviewed as required. Any problems identified are discussed with leads to have them rectified. If any services were found to be non-compliant, a process is in place to escalate to the commissioning officers within the ADP. With regards to ACA and Addiction Waiting Times data for the NHS Board, reports are available for them to run prior to the monthly upload. In addition, a number of checks/monitors are carried out in the individual services themselves to ensure the data is as accurate as possible.

Every new member of staff receives one-to-one training on how to complete the waiting times section as well as information on calculating the referral to treatment waiting times, clock stopping for DNA and cancelled appointments by the service user and the appropriate use of 'periods of unavailability'. NHS staff also have to pass additional training. Each locality within the NHS has their own waiting times champion, a member of staff with increased knowledge of access to treatment targets who supports other staff in provision of treatment to service users.

A compliance checklist is kept at ISD to help monitor the return of the compliance forms. Compliance is high. The compliance tables are published quarterly by ADPs. There is a section called 'service compliance' within each Drug & Alcohol Waiting Times publication, within which there are links to the compliance data. A separate table exist for prison data by ADP.

ISD Data Management interrogates the returns to check for anything unusual in the returns as an added layer of quality assurance and send out reports on a monthly basis. ISD produces tables on the completeness of data collection on a quarterly basis.

The public interest profile of the statistics is low, as reflected by low media interest and reporting against a national standard. The level of risk of quality concerns, given that manual entry and thorough checking processes exist, is medium to high. As a result, ISD is of the opinion that the DATWT statistics meet level A2 enhanced assurance in relation to the operational context and administrative data collection requirements, communication with data supply partners and QA principles, standards and checks applied by data suppliers, outlined in the UKSA data quality assurance matrix.

3.4 National Records of Scotland data

3.4.1 Populations

National Records of Scotland (NRS) produce population data as at 30th June each year. The UK Statistics Authority undertook an assessment of the population estimates and projections in 2015. Further information is contained in the link below:
Population estimates and projections for Scotland (National Records of Scotland)

NRS provides further information about the quality assurance of the population data on their website:
Information on the quality assurance arrangements for administrative data used in population estimates

The population data used by the ISD analysts is taken directly from the NRS website. As this data is publicly available on the website it is not supplied directly to ISD. Therefore, there is no service level agreement in place. Nevertheless, additional checks are run to ensure the data is extracted correctly, by checking the ISD file against the published results on the NRS website.

The Population and Migration Statistics Group (PAMS) meets every 6 months, and there are representatives present from NRS, ISD and local authorities. During these meetings any issues with the data are discussed.

The data are of medium quality concern and medium public interest. Also, ISD has confidence in the the quality of the UKSA assessment that the data are subject to as well as the additional checks carried out by ISD. ISD is therefore of the opinion that the population datasets meet level A2 enhanced assurance.

3.4.2 Deaths

Deaths occurring in Scotland are registered by a network of Registrars located throughout Scotland, and the records are collated centrally by the National Records of Scotland. These data are then used by ISD for a very wide range of analyses, including those for the Scottish Public Health Observatory ScotPHO topics of Suicide and Healthy life expectancy. ScotPHO is a collaboration co-led by ISD Scotland and NHS Health Scotland, and includes the Glasgow Centre for Population Health, National Records of Scotland and Health Protection Scotland.

Provisional death data is sent to ISD on a weekly basis from NRS and the final data issued around mid August each year. NRS provide information on some specific checks of the data, the quality of some of the items about deaths held in the vital events statistical database, and the comparison of NRS and NHS data. Visit the following site for further information:
http://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/deaths/deaths-background-information/quality-of-nrs-data-on-deaths

In addition to the processes already carried out by NRS, ISD will run certain checks, if relevant for the publication, to ensure that the records received by ISD are fit for purpose. Regular SMR Analyst Forum (SAF) meetings take place every 6-weeks and membership includes ISD analysts and data management professionals. The NRS Death data is on the agenda for discussion.

As the ISD ScotPHO suicide publication is complimentary to the NRS suicide publication, and is released on the same day, the two organisations liaise about their plans, including how to handle changes such as suicide coding and the European Standard Population for standardising. The two organisations send pre-release text and tables to each other for quality assurance purposes.

Given the safeguards in place including ISD's own checks and UKSA assessment of NRS statistics, ISD is of the opinion the death data meet level A2 enhanced assurance. The deaths data have medium quality concern and medium public interest.

3.5 Self-Assessed Health (SAH)

The following data are required to calculate Life Expectancy (LE) and Healthy Life Expectancy (HLE):

- Deaths data
- Population estimates data
- Self Assessed Health (SAH) data

This section will focus on the Self Assessed Health (SAH) data, as the NRS Death and Population data have been covered in Section 3.4.
Data for Self Assessed Health (SAH) are based on two surveys:

- the Scottish Health Survey (SHeS) and
- the Scottish Household Survey (SHoS).

The UKSA undertook an assessment of the SHoS in December 2009:
https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/12/images_assessment-report-26-assessment-of-scottish-household-survey-outputs_tcm97-28970.pdf

The response to the UK Statistics Authority provides further details, in particular Section C on Quality Assurance Standards:
http://www.gov.scot/Resource/Doc/933/0113252.pdf

The following link provides further information on the Scottish Health Survey analytical quality data control:
http://www.gov.scot/Topics/Statistics/Browse/Health/scottish-health-survey/SHeSDataQualityControl

Data from SHoS are sent to ISD from the Scottish Government. The SHeS data are also received from the Scottish Government. There are more formal procedures in place for the SHeS data, and an audit trail is kept from receipt until it is destroyed, which was in the terms of agreement.

The self assessed health (SAH) files are checked using the excel auto filter facility to check for any anomalies such as missing values or negative values etc.

Further checks are made once the data has been imported in SPSS to check for the validity of the data.

Regular meetings take place between ISD, NRS and Office of National Statistics (ONS).

These household survey statistics have high quality concerns and low public interest. ISD is of the opinion that the datasets feeding into the HLE publication meet level A1 basic assurance. ISD relies on an intermediary data supplier. However ISD is satisfied the SHoS

undergoes regular assessments by the UKSA. The Scottish Government has publicly made information available on data quality control in relation to SHeS data.

3.6 Further Checks

The analysts carry out additional checks to ensure the data is fit for purpose before analysis. There is regular engagement between data management teams and the analyst teams responsible for producing the publications in the form of emails and meetings.

The checking guidance internal document for ISD provides staff producing analytical outputs with a reference for best practice for checking data analysis. The principles apply to all analytical/statistical work including publications, project work, information requests (including Freedom of Information requests) and parliamentary questions across all datasets. By following these best practice principles and using a checklist it ensures that ISD takes a consistent approach to checking all data analysis. An extract of the checklist is shown in Appendix 2.

4. Quality Assurance of Outputs

In addition to the checks of syntax, methodology, tables, visualised outputs carried out locally by ISD analytical teams, there are three layers of scrutiny undertaken within ISD, which involves senior staff participation, before the publication of official statistics:

- Content Review Board
- Editorial Board
- Publications Briefing

Content Review Board

The aim of the Content Review Board is to positively inform and influence the way in which ISD and Health Protection Scotland (HPS) release content to its customers. Improvements to the content of the ISD and HPS outputs are considered, focusing on the following aspects:

- Developing 'information into intelligence':
    - Increasing the statistical content of the publications
    - Increasing the health inequalities content
    - Increasing the degree of public health content
    - Increasing the clinical interpretation
    - Increasing relevant contextual information
- Improving the structure, organisation and language of outputs;
- Developing good practice in the use of 'plain English' and increase its use across publications;
- Identifying where content could be released in a different form than at present;
- Increasing the benchmarking content of releases;
- Combining or aligning releases in order to provide a wider overview of a subject area.

The content review board meetings take place every 8-10 weeks, although the meetings can take place more frequently. The core group membership comprises the clinical director, epidemiologist, communications manager and service managers. Additional volunteers from within ISD provide additional support, e.g. analysts.

The Content Review Board subjects the publication to an intense level of scrutiny, including focusing on how well the completeness and quality of the data are explained in the publication and that any tables are footnoted clearly and correctly, along with any supporting Excel files. The Board also advises on the methodological approach and use of data visualisation for increasing the value of the outputs.

Editorial Board

The Editorial Board planning meeting has been established to have an oversight of all the official statistics publications to ensure there is a consistent approach. The group meets every month to discuss publications going out in the following two to three months. The Board comprises a selection from the Associate Director, Heads of Service, Consultants in Public Health Medicine, and the communications manager.

The authors of the publication are invited, along with the Service Manager for the area, to present their plans for the forthcoming publication. In particular, the focus is on any factors that have changed since last time that will affect what is published (e.g. policy changes, data collection changes), any improvements intended for the publication (e.g. context, analysis, interpretation, presentation of results and dissemination methods), how the improvements will be carried out (e.g. types of consultation, stakeholder engagement, internal input), how the publication links in with any other publication (e.g. within ISD, National Services Scotland, the Scottish Government, or UK-wide), and plans to include data visualisation.

Each publication will be allocated a 'sponsor' from the Editorial Board Planning Meeting group members, who will be available to support the analysts making any changes necessary and help review the final output, discuss the changes that have been made and the presentation of the final report. Any publication which has altered substantially since the previous version may require the updated version to be presented to the Editorial Board.

Where visualisations are being considered, analysts must include plans to consult with a member of the ISD's Data Visualisation Group who will work with the analysts to develop ideas and advise on good practice.

Key Messages Handling Meetings

There are weekly internal handling meetings held every Thursday within ISD. These consist of the associate directors, heads of service, the consultant in public health medicine responsible for information, service managers and the communications manager. During these meetings, the analytical teams whose statistics publication will be available for pre-release access the following Tuesday are invited to present the publication summary (which includes the key messages) and agree on the messages and infographic for ISD's Twitter feed. The meeting participants are briefed on any key findings that are likely to cuase considerable media interest on release. The communications manager ensures that plain English is used throughout the publication summary and it may be necessary to work with

the analyst at the end of the meeting to act on the recommendations from the meeting participants.

5. Revisions Policy

Scheduled revisions are managed in accordance with ISD Revisions Policy:
http://www.isdscotland.org/About-ISD/About-Our-Statistics/ISD-Revisions-Policy-V04.pdf

6. User Engagement

In the Code of Practice for Official Statistics, Principle 1, Practice 1 states that producers must engage effectively with users of statistics to promote trust and maximise public value, in accordance with *Protocol 1: User Engagement*.

ISD is keen to seek the views of users of health statistics in Scotland to improve the quality, value, accessibility and impact of its outputs. A joint engagement event in Edinburgh was arranged in 2014 with ISD, the UK Statistics Authority and health statistics users throughout the UK.

There is ongoing user engagement throughout ISD. To aid this process, ISD and HPS staff use a consultation guidance document to assist them in selecting the most appropriate consultation method.

On 12 April 2016, ISD published the Prescribing Activity Official Statistics in new formats – open data and Tableau enabled interactive data visualisation. It was supported by consultation with users that took place from September to October 2015. The consultation document and metadata are published on the ISD website.

In July 2016, ISD's secondary care team carried out a consultation with users which resulted in a new way to present their acute activity and hospital beds publications.

There have been a number of consultations with users by ISD's waiting times teams to ensure users are engaged when there are plans to change the format, timeliness and content of outputs. ISD plans to go through a full consultation process with all users as part of a plan to combine a number of waiting times publications, including all of the key Scottish Government and NHS Board contacts, as well as inviting feedback from the public through release of the consultation on the ISD website.

ISD's smart publishing improvement programme is currently working with analytical teams to develop new ways of presenting and disseminating their statistics in order to engage with a wider range of users add value to statistical outputs. User consultation documents are being prepared.

There are many regular meetings between data management staff, analytical teams and data suppliers (e.g. NHS Boards) to discuss matters pertaining to quality of data at source. The feedback helps improve the quality of sources, inputs and outputs.

A link is contained within every publication to give users the opportunity to provide feedback.

Appendix 1: Administrative Data Quality Assurance Matrix – UKSA Recommendations

Level of Assurance	Areas of practice related to quality assurance of administrative data regularly provided for producing Official Statistics			
	Operational context & administrative data collection	**Communication with data supply partners**	**QA principles, standards and checks applied by data suppliers**	**Producer's QA investigations & documentation**
A0: No assurance	• Context and data collection by supplier not investigated, managed or documented	• No communication	• No description of suppliers' QA procedures and standards	• No description of own QA checks
A1: Basic assurance Statistical producer has reviewed and published a summary of the administrative data QA arrangements	Consider the following types of activities: • Producer has provided users with an outline of the administrative data collection process, • Illustrated the administrative data collection process and main stages, • Outlined the operational context, • Identified actions taken to minimise risks to quality, • Identified and summarised the implications for accuracy and quality of data, including the impact of any changes in the context or data collection	Consider the following types of activities: • Producer has outlined the data provision arrangements including: - annual statement of needs, - timing and format of data supply, - coordination of data sign-off from supplier, • Fed back identified errors to data suppliers and recorded their response, • Sought the views of statistics users about the data and resolved any quality issues reported	Consider the following types of activities: • Producer has knowledge of suppliers' QA checks and published a brief description, • Identified whether audits are conducted on the admin data (such as internal or operational audits, external audit such as by regulator), • Described the implications for the statistics	Consider the following types of activities: • Producer has established regular QA checks on the received admin data, • Published a description of its own QA checks on the admin data, • Outlined general approach and findings, • Identified the strengths and limitations of the admin data, • Explained the likely degree of risk to the quality of the admin data
A2: Enhanced assurance Statistical producer has evaluated the administrative data QA arrangements and published a fuller description of the assurance	Consider the following types of activities: • Producer has provided users with a fuller description of the operational context and administrative data collection arrangements, eg: - a process map detailing data collection processes, - explanations for classifications, • Identified and summarised potential sources of bias and error in administrative system, • Identified and described safeguards taken to minimise risks to data quality, • Provided a detailed description of the implications for accuracy and quality of data, including the impact of any changes in the context or collection arrangements	Consider the following types of activities: • Producer has agreed and documented: - data requirements for statistical purposes, - legal basis for data supply, - data transfer process, - arrangements for data protection, - sign-off arrangements by data suppliers, • Established an effective mode of communication with contacts (eg with data collector and supplier bodies, IT systems, operational/policy officials) to discuss the ongoing statistical needs in the data collection system and quality of supplied data, • Sought the views/experiences of statistics users and resolved any quality issues reported	Consider the following types of activities: • Producer has provided a fuller description of the main QA principles, quality indicators and checks used by the data suppliers, • Described the role of relevant information management or governance groups in data quality management, • Described the role of audit of the admin data within the collection and operational settings, • Described the implications for the statistics for the quality issues identified by data supply bodies and regulators	Producer has provided a fuller description of its own QA checks on the admin data, • Detailed the general approach and findings for specific quality indicators, • Identified the strengths and limitations of the admin data, • Explained the likely degree of risk to the quality of the admin data
A3: Comprehensive assurance Statistical producer has investigated the administrative data QA arrangements, identified the results of independent audit, and published detailed documentation about the assurance and audit	Consider the following types of activities: • Producer has provided users with a detailed description of the administrative system and operational context: - explained why the data are collected, who by and how, - identified differences across areas in the collection and recording of the data, - identified issues for individual data items, such as whether objective or based on subjective recording, missing and/or imputed, poorly recorded, • Identified issues in design and definition of performance measurements and targets, • Identified and described potential sources of bias and error in the administrative system, • Identified and explained any safeguards used to minimise the risks to data quality, • Provided detailed description of the implications for a quality of the data, including the impact of any changes in the context or data collection	Consider the following types of activities: • Producer has established/maintained collaborative relationships, • Has a written agreement specifying: - roles and responsibilities, - legal basis for data supply, - data supply and transfer process, - security and confidentiality protection, - schedule for data provision, - content specification, • Used a change management process, • Regularly communicated with the data collector and supplier bodies, IT systems, operational/policy officials eg newsletters, conferences, attending data supplier/IT system group meetings, • Regularly engaged statistics users, resolved any reported quality issues, and held user group conferences	Consider the following types of activities: • Producer has described the data suppliers' principles, standards (quality indicators) and quality checks, • Reviewed quality reports for the received data (such as input quality indicators for data accuracy, coverage and completeness), • Identified and documented the findings of investigations and audits conducted on the admin data and associated targets (such as internal and operational audits, and external audits by regulators and professional bodies), • Described the implications for the statistics and determined whether the data continue to be satisfactory for official statistics purposes	Consider the following types of activities: • Producer has provided a detailed description of its own QA checks on the admin data (including validation, sense and consistency checks), • Given quantitative (and where appropriate qualitative) metrics for specific quality indicators (such as input, process and output quality metrics), • Undertaken comparisons with other relevant data sources (such as survey or other admin data), • Identified possible distortive effects of performance measurements and targets, • Identified the strengths and limitations of the admin data and any constraints on use for producing statistics, • Explained the likely degree of risk to the quality of the admin data

Appendix 2: Summary Checklists

A2.1 First Person Check

Before you start your analysis you should:

- ☐ Know/find out who is responsible for the dataset(s) you are analysing (the 'owner')
- ☐ Check whether the same/similar analysis has been produced before
- ☐ Decide on required documentation (commented program, full SOP etc)
- ☐ Discuss coding choices with the relevant clinical expert.
- ☐ Check for data completeness issues
- ☐ Check for data quality issues, referring to established metadata where available
- ☐ Check which is the most recent time period published in any Official Statistics release and therefore available for use for information request/PQ responses

When checking your outputs you should:
A: check your syntax and data selections

- ☐ Check the programme/query is doing what you expected
- ☐ Check data are being read in accurately
- ☐ Check you are using the latest file layouts
- ☐ Check you are selecting the correct codes
- ☐ Check you have selected all the required fields/inclusions/exclusions

B: perform basic logic checks of the output

- ☐ Check that figures are what you would expect
- ☐ Check that comparisons between the Boards look reasonable

Remember – assume outliers or unusual patterns should be considered to be a data quality issue until thorough investigation has eliminated this possibility

- ☐ Check that subtotals add up to totals
- ☐ Check if suppression is required for Disclosure Control purposes

C: compare your outputs with other sources

- ☐ Check your outputs against other publications/outputs (PHI or external) that contain the same or related information.

D: rerun the analysis

- ☐ Check that an alternate analytical package gives the same result
  Or
- ☐ Check that the same program/analysis gives accurate results on a set of known data (e.g. previous year)

E: check data transfers

- ☐ Check that data has been transferred accurately between software packages (e.g. from BOXI to excel)
- ☐ Check formulae are accurate (e.g. in excel calculated cells)
- ☐ Check formulae that have been 'dragged' are accurate for all cells

- ☐ Check that NHS Board descriptions and data are in the same order
- ☐ Check that any 'sorts' have included the appropriate data array
- ☐ Check that figures quoted in narrative sections are accurate

F: formatting and presentation

- ☐ Check all titles are correct
- ☐ Check all data sources are listed
- ☐ Check all caveats and footnotes are added and accurate
- ☐ Check spelling
- ☐ Check for overall understanding and clarity of the output
- ☐ Check any data visualisations with the Data Visualisation Core Group
- ☐ Check for accessibility
- ☐ Check print previews to ensure page settings are appropriate
- ☐ Check excel sheets open on the required sheet, curser on cell A1
- ☐ Check formatting is consistent across all tables/documents/charts

After analysis is complete you should:

- ☐ Document the analysis carried out/comment within programmes/syntax
- ☐ Update Standard Operating Procedures if required
- ☐ Apply the Statistical Disclosure Control Policy
- ☐ Ensure the outputs meet the requirements specified in any Public Benefit & Privacy Panel (PBPP) authorisation, if applicable

A2.2 Second Person Check

The range and depth of checks carried out by the second person will be determined by the particular analysis being produced, local team protocols and your Service/Group Operations Manager but you should consider:

- ☐ Check the final output matches the specification of request.
- ☐ Check syntax/method
- ☐ Sense check of figures
- ☐ Check analysis against other sources / previous analysis / publications
- ☐ Check by rerunning high level analysis in another package
- ☐ Check transfer of figures from output to a different package
- ☐ Format checks on final Table/Chart.
- ☐ Check supporting correspondence to customer
- ☐ Check that there is a completed statistical disclosure control assessment
- ☐ Check that the outputs meet the requirements specified in any PBPP authorisation, if applicable