Office for Statistics Regulation

Strengthening the quality of HMRC official statistics

April 2020

Office for Statistics Regulation

We provide independent regulation of all official statistics produced in the UK. Statistics are an essential public asset. We aim to enhance public confidence in the trustworthiness, quality and value of statistics produced by government.

We do this by setting the standards they must meet in the Code of Practice for Statistics. We ensure that producers of government statistics uphold these standards by conducting assessments against the Code. Those which meet the standards are given National Statistics status, indicating that they meet the highest standards of trustworthiness, quality and value. We also report publicly on system-wide issues and on the way statistics are being used, celebrating when the standards are upheld and challenging publicly when they are not.

Acknowledgements

We would like to express our gratitude to everyone who gave their time to speak with us, for their cooperation and openness, to inform this review.

Contents

Foreword

This is the first time we have looked at how the quality of official statistics is managed across a whole department. We have found much to admire in the way that HM Revenue and Customs (HMRC) produces official statistics, but several areas where it could strengthen its approach: we make a series of recommendations to address these issues.

This is an important report not just for HMRC but for all statistics producers. While the report highlights recommendations for HMRC, I would like to emphasise broader points that emerge:

- **Analysts do not operate in isolation** The wider organisational context in which statistical producers work will always impact their ability to produce statistics that are of appropriate quality. To facilitate effective quality management, organisations need to have well-designed and managed systems and processes, well-established relationships between partners, and actively promote consistent quality standards and values. Quality is everyone's responsibility.

- **Often the biggest risk comes at the beginning of the production process** There is a big risk to the quality of statistics when analysts do not fully understand the nature and quality of administrative data they are working with. It is really important that all producers assure themselves of the quality of their data and build effective working relationships with data providers or managers to facilitate this. This is especially true if working with older IT systems that have typically been built without analytical requirements in mind, which can create additional obstacles for effective data extraction and analysis.

- **Quality assurance could be strengthened** We expect all producers to be considering a broad approach to quality assurance, such as: developing their understanding of the data collection and preparation process; peer reviewing and carrying out deep dives of their analyses; and comparing their statistics to independent data sources. Producers across the GSS are starting to demonstrate the positive role programming languages like R or Python can have in facilitating quality assurance and management.

- **Producers of statistics will never eliminate errors entirely** Producers should strive to establish and maintain effective processes and policies that lower the risk of errors, especially errors that have a large impact on the interpretation of the statistics and therefore for users.

- **Sometimes it's better to do less** Demands on analysts are often high. We'd encourage all producers to review, in consultation with users, their current publications, with a view to reducing the number or size. This can free up resource and, at the same time, help producers to fulfil their commitment to producing statistics of public value that meet user needs.

- **Strong leadership and knowledge networks enable a strong quality culture** It is really important that analytical leaders champion and support changes and developments that can enhance quality. This gives statistical producers permission to improve and innovate. Prioritising knowledge sharing about quality across areas working with data and statistics, and building effective mechanisms to do this, will strengthen a culture that has quality embedded at its core.

By inviting this review, HMRC has taken a proactive and open approach to strengthening the quality of its official statistics. Our findings will ring true for other Departments and organisations in Government, and I'd encourage other producers to reflect on the recommendations and to think about how they can build on their own approach to quality, to ensure statistics meet the needs of the people who use them.

1. An Introduction to the review

In September 2019, HMRC invited the Office for Statistics Regulation to carry out a review of the principles and processes underpinning the quality of HMRC's official statistics. This review was proactively initiated after HMRC identified a significant error in published Corporation Tax receipt statistics, which affected the period from April 2011 to July 2019.

Aim and scope of the review

The aim of our review was to provide an independent assessment of the approach that HMRC takes to manage quality and risk in the production of its official statistics and to identify potential improvements. We appreciate that producers of statistics will never eliminate errors entirely: the recommendations we present in this report focus on improvements that HMRC should make to help minimise the risk of issues with its statistics in the future.

While the error in published Corporation Tax statistics was an important trigger for this review, the review has not probed into this issue specifically: HMRC has carried out a range of other pieces of work, including an internal audit of the circumstances around the Corporation Tax error, to fully evaluate this issue and to inform corrective actions and improvements to mitigate the risk of a similar issue in the future. Corporation Tax estimates for the period were corrected by HMRC in its 'Tax and NIC receipts: information and analysis (August 2019)' publication on 24 September 2019. The scale of the revisions varies between years, from the revised receipts figures being £0.6bn lower than previously published in 2011/12 through to £4bn lower in 2018/19. More information about the error is given in Annex 1: HMRC Corporation Tax Statistics.

We have not undertaken detailed investigation of the statistical methods used to produce HMRC published statistics. Our focus has been on the processes and principles underpinning the production and quality assurance of these statistics, which aim to mitigate risk and avoid issues, including errors, with the statistics.

The Office for Statistics Regulation

The Office for Statistics Regulation is the regulatory arm of the UK Statistics Authority. We provide independent regulation of all official statistics produced in the UK.

Statistics are an essential public asset and we aim to enhance public confidence in the statistics produced by government. In all our work, we draw on our Code of Practice for Statistics.

The Code encompasses three pillars: Trustworthiness, Quality and Value. Each pillar contains a number of principles and detailed practices that producers of official statistics commit to when producing and releasing official statistics. Complying with the Code gives users of statistics and citizens confidence that published government statistics are of public value, are high quality and are produced by people and organisations that are worthy of trust. During this review, we have considered whether HMRC is fulfilling its commitments to users and citizens, as outlined in the Code, and we refer back to the Code to anchor our recommendations in the sections below.

Our approach

HMRC produces in the order of 75 official statistics, many of which are published on several occasions throughout the year. These official statistics are almost exclusively based on administrative data from HMRC's own IT systems. HMRC's Knowledge, Analysis and Intelligence (KAI) directorate is responsible for around 65 of these official statistics, as well as a broad range of other analytical outputs, including those used internally within HMRC. Around one-third of the 450 analysts working in KAI contribute to official statistics production: for most of these analysts, the production of official statistics is only part of their wider role.

More information about the scale of HMRC as an official statistics producer, including the structure of KAI and a list of HMRC publications, is included in Annex 2: HMRC as a producer. The OSR review focused on four of the five business areas in KAI that produce official statistics: Benefits & Credits, Direct Business Taxes, Indirect Taxes Customs & Co-ordination, and Personal Taxes. These four areas produce the vast majority of HMRC's official statistics. The fifth area, Customer Compliance and Strategy, produces one official statistic.

This review considered:

- the overall approach to quality management in KAI and the roles and responsibilities of statistical producers at all levels, as well as their managers.
- existing quality assurance (QA) processes and current improvement plans for official statistics within KAI.
- the processes in place for managing risk and handling errors in statistical production.
- the levels of risk of different sources of error in KAI's production processes.
- the information about quality that KAI makes available to users of its official statistics.

To be able to make tangible recommendations about how HMRC can improve the way it produces official statistics, and minimise the risk of errors, we:

- interviewed a wide range of people who produce or oversee the production of official statistics in KAI. This helped us to understand how the quality of statistical outputs is currently managed and how quality assurance is carried out in HMRC.
- interviewed individuals from several HMRC teams outside of KAI, who provide data to KAI, which is used in HMRC's official statistics. To produce official statistics, KAI needs to work in partnership with other parts of HMRC, so any recommendations made by us had to consider its broader context. Therefore, we were interested in the arrangements that KAI has in place to meet its data requirements and how KAI works with those who provide it with data.
- interviewed a selection of external stakeholders, who are key users of HMRC statistics. Although not the primary focus of the review, we considered how HMRC interacts with key users who rely on its statistics and its transparency about issues with these statistics.
- reviewed existing quality management and assurance guidance used within KAI.
- reviewed guidance relating to quality from the Government Statistical Service (GSS) Best Practice and Impact Division and looked at how other departments manage quality. This was to identify examples of good practice to help shape our recommendations to HMRC.

A full list of those we engaged with during the review is provided in Annex 3: List of meetings undertaken during this review. We would like to express our gratitude to everyone who gave their time to speak with us, for their cooperation and openness.

Audience for this report

The immediate target audience of this review is HMRC official statistics producers at all levels, as well as their managers, up to Permanent Secretary and ministerial level. Although this review focussed on HMRC's production of official statistics, the findings and recommendations contained within this report are relevant to analysis within KAI more generally and will be applicable to a wide range of teams in HMRC, outside of KAI, who collect, manage and analyse data. Some of the

findings and recommendations also apply directly to operational areas of HMRC, such as Finance and IT: these functions have an important role in enabling KAI to produce quality assured statistics.

Beyond HMRC, we hope that publishing the outcomes of this review will benefit other statistical producers within the GSS and encourage them to review their own quality management and assurance processes.

Finally, we are confident this report will help HMRC achieve its ambition to be transparent with users of its statistics about the issues it has faced, and how it is addressing them going forward.

Next steps

HMRC will provide a response to this review, which explains to users of HMRC statistics the actions it will take to improve quality management and assurance of its official statistics. OSR will consider how future regulatory work that focusses on HMRC statistics can be used to challenge and support the actions HMRC is taking to improve quality in its statistics.

2. Executive summary

Organisations that effectively manage the quality of the data they produce and use have well-designed and managed systems and processes, well-established relationships between partners, and actively promote consistent quality standards and values. The operation and credibility of any statistical organisation is risked when quality management of data is not prioritised.

HMRC produces lots of official statistics and most of the time it does this well

Data are vital for the functioning of HMRC: its core activities rely on high quality and accessible data and statistics. HMRC produces many official statistics and the demands on analysts in KAI are high: most analysts who produce official statistics do so as part of a wider set of responsibilities. It is important to recognise that HMRC produces the majority of its official statistics regularly without issues or errors. Indeed, in the course of our review, we met teams who are managing the quality of their statistics very well. Throughout this report, we present examples of good practice we have heard about, which serve to illustrate the good work already happening within KAI that others can learn from and build on. KAI is currently updating its existing quality assurance frameworks, which demonstrates its ambition and progress in this area.

However, quality assurance and quality management could be strengthened

We have seen several common themes and areas for improvement emerge from our review. At present, quality assurance is not always adequate across teams, and initiatives that would support improvements in quality management across KAI are not always prioritised. KAI needs to create and maintain a culture that invests more in quality, which will require commitment and action from producers at all levels, including senior managers, as well as an assessment of the quantity and allocation of available resource in KAI. If KAI does not create this culture, the production of statistics could resemble a series of mechanical processes that do not always identify significant quality issues. KAI does not operate in isolation from other parts of HMRC and long-term improvements to quality management will rely on understanding, values and responsibility being shared across many areas of HMRC, especially those that own or process administrative data.

We have nine recommendations to strengthen the quality culture in HMRC

We have nine recommendations, explained in detail in the *Findings and recommendations* section of this report. Implementing these recommendations will help strengthen the quality culture in HMRC and make sure that the structure and tools are in place to manage statistical quality effectively. This will minimise risks to the quality of HMRC official statistics. The most immediate of our recommendations relate to establishing and maintaining effective relationships between KAI and other parts of HMRC it receives data from; understanding data journeys; and quality assurance of data, analysis and publications. Further recommendations, which will help enable the immediate recommendations, focus on strengthening the leadership and oversight of quality management in KAI; and building more effective routes for knowledge sharing about quality within KAI. Two of our recommendations relate to improving the information KAI publishes about the quality of its official statistics, which would enhance the trustworthiness of HMRC as a statistics producer. A table summarising our findings and recommendations is included in section 3 and repeated in Annex 4: Key findings and recommendations of the review.

HMRC will need to invest resource into making and sustaining necessary changes

Making, embedding and managing the changes outlined in this report will require resource. HMRC should consider the changes that need to be implemented and come to its own conclusions about whether additional resource will be needed in KAI or the areas of HMRC who work with KAI, or if the changes can be achieved by using existing resources differently. In either case, there may be an up-front requirement for additional resource to review and change current practices. More widely, further investment in HMRC's IT systems would be beneficial to ensure they keep pace with modern expectations of functionality and flexibility, to better serve the needs of data analysis.

3. Key findings and recommendations of the review

The table below summarises the key findings and recommendations of this review of the principles and processes underpinning the quality of HMRC's published official statistics. More detail about the findings and recommendations can be found in section 2 and section 4 of the report.

Table 1: Key findings and recommendations

	Finding	Recommendations
Leadership and oversight of quality management in KAI		
1	Senior leaders in KAI do not always give sufficient priority and resource to improving quality management and quality assurance. (Finding 4.1.4)	Senior leaders in KAI should champion and support developments and innovations that can enhance the quality of official statistics. (Recommendation 4.1.7)
2	There is little opportunity for routine sharing of good practice with respect to quality, or learning from past issues in statistics, across KAI. (Finding 4.1.6)	Knowledge sharing should be prioritised at all levels of KAI, and more-effective mechanisms built to support this. As part of this, a cross-KAI quality network could be put in place, with senior support. (Recommendations 4.1.8 to 4.1.9)
Working with data providers and QA of input data		
3	Producers in KAI do not always have effective relationships with data suppliers, many of whom are from other parts of HMRC, and do not always fully understand the nature and quality of the data they receive. (Finding 4.2.1)	In some areas, there is the need for closer working with teams providing data to KAI. The responsibility for establishing and maintaining effective relationships should be shared between KAI and its partners. (Recommendations 4.2.5 to 4.2.6)
QA of analysis and publications		
4	Not all teams consider the whole data journey, from the initial data input to a statistical publication when carrying out QA. (Finding 4.3.1)	Each team should produce and maintain a process map, which shows how its statistics are produced, from data collection to publication. This map should highlight risk points and mitigating actions taken to ensure the quality of the statistics. (Recommendation 4.3.5)
5	The application and completeness of QA done during the statistical production process is not always adequate across teams. (Finding 4.3.1 to 4.3.3)	There should be a more consistent approach to, and clearer expectations for, QA across KAI. QA should be documented in each production round and available for audit. (Recommendations 4.3.6 and 4.3.7)
6	More teams could be using data science methods and tools to effectively manage and peer review code, remove error-prone steps in production and create efficiencies. (Finding 4.3.4)	Drawing on its relationship with the HMRC Data Exploitation team, KAI should move towards greater use of data science methods and tools. (Recommendations 4.3.8 and 4.3.9)
Published materials relating to quality of HMRC official statistics		
7	Much of the published information on quality of HMRC official statistics does not fully explain to users the strengths and limitations of the statistics and data, or the production process. In some cases, it does not appear to have been updated since 2013. (Finding 4.4.1)	HMRC should review the published information about the quality of its official statistics and update to ensure that it fully explains how HMRC producers assure themselves that their statistics are accurate, reliable, coherent and timely. (Recommendations 4.4.3 and 4.4.4)
8	Published information about issues with and delays to official statistics can lack detail and is not easy to find. (Finding 4.4.2)	HMRC should publish information in a transparent way that is easy to find. (Recommendation 4.4.5)
Analytical resources and training		
9	Analysts producing statistics in KAI can struggle to find time to invest in developing and improving their statistics. Analytical resource that could be used for this may be being spent on producing publications with more content, or that are produced more frequently, than needed by users. (Finding 4.5.1)	Each area of KAI should review its current publications, with a view to reducing the number or size of publications, to ensure analytical resource is being used effectively. (Recommendation 4.5.6)

4. Findings and recommendations

This section of the report presents a discussion of our findings from the review and the areas that HMRC should address to improve quality management and assurance of its official statistics. The responsibility for producing high quality statistics lies with everyone working in the statistical production system: the following discussion provides areas for consideration for people working at all levels in KAI, as well as for those who oversee the work of KAI and for wider HMRC. The Code of Practice for Statistics sets out the practices that HMRC is committed to in its production of official statistics: HMRC should use the Code to help deliver its ambition to improve the quality management of its statistics and data.

4.1. Leadership and oversight of quality management in KAI

All organisations producing official statistics should be open about their commitment to quality and about their approach to quality management. They should ensure that organisational structures and tools are in place to manage quality effectively and promote appropriate quality standards.

To meet this principle, and thereby to enable effective management of risks to quality, the roles and responsibilities of all those involved in the production of statistics should be clearly defined. Managers who provide clear expectations with respect to quality, as well as guidance and support, will better enable junior producers to understand and carry out their roles, minimising the risk of error. Senior leaders should prioritise quality management and promote and support developments and innovations that can enhance the quality of official statistics. They also need to act as advocates for official statistics to ensure that they are valued across their organisation.

Findings on leadership and oversight of quality management in KAI

4.1.1. Several aspects of KAI's approach to producing statistics support effective quality management. The HMRC Head of Profession for Statistics (HoP) sits within KAI at a senior level and has a Deputy HoP. Those working within KAI feel well supported by the HoP and Deputy HoP on matters relating to the Code of Practice and feel that they can approach them for help if they have any issues with their statistics. The Deputy HoP maintains a log of issues that have occurred in the production of HMRC official statistics, which is reviewed with the HoP to help identify areas of high risk and to guide actions to prevent systemic problems emerging. Significant issues are flagged with the Director of KAI.

4.1.2. The HoP and Deputy HoP are supported by a 'Lead Statistician' in each of the four areas within KAI that produce the bulk of HMRC's official statistics. Lead Statisticians act as a first point of call for questions relating to official statistics production and the Code of Practice in their area, which offers teams producing statistics another level of support.

4.1.3. There is a consistent framework (outlined in Annex 2: HMRC as a producer) across KAI for roles and responsibilities of staff at different levels in KAI, with respect to quality management in the production of individual official statistics outputs – from obtaining raw data and analysis, through quality assurance, to final sign off. This framework has come from the HoP and staff at all levels in KAI are clear on it.

4.1.4. Although there are clear roles and responsibilities for individuals with respect to quality when producing individual official statistics, against a backdrop of a range of other, sometimes conflicting, high-profile demands, senior leaders do not always give sufficient priority and resource to improving quality management and quality assurance. This includes setting an expectation for really full, comprehensively-documented quality assurance of analysis, encouraging those producing statistics in KAI to dedicate resource to initiatives that support quality beyond their own statistics (such as investing time in sharing good

practice and learning) and providing structured leadership and support for innovation, such as using data science methods and tools.

4.1.5. In some cases, senior leaders may have clear plans and ambitions for official statistics, but these are not always visible to analysts working within KAI. For example, while the HMRC HoP and Deputy HoP have an improvement plan for official statistics in KAI, which sets out the ambition to roll out data science tools and capabilities across KAI, for the most part, the people we spoke to who produce official statistics in KAI were not aware of this plan.

4.1.6. At present, there is not an effective route for routinely discussing or disseminating information about quality across KAI, nor to strengthen a culture that has quality embedded at its core. Although there is an HMRC quality champion, who is part of the cross-Government Statistical Service (GSS) quality network, this role is under-resourced.

Recommendations on leadership and oversight of quality management in KAI

Recommendation 1: Senior leaders in KAI should champion and support developments and innovations that can enhance the quality of official statistics.

4.1.7. A more visible KAI-wide strategy for improving and managing quality, which comes from and is championed by senior leaders, would help teams producing statistics to know what level of quality is required of the statistics. Senior leaders should proactively champion and support developments and innovations that can enhance the quality of official statistics, to ensure they can happen and be sustainable.

Recommendation 2: Knowledge sharing should be prioritised at all levels of KAI, and more effective mechanisms built to support this. As part of this, a cross-KAI quality network could be put in place, with senior support.

4.1.8. To enable more-effective communication about quality among teams and areas, KAI needs to prioritise knowledge sharing and build more effective mechanisms to do this. Better knowledge sharing across teams and areas, at all levels of KAI, is needed to raise awareness among producers in KAI of the steps they can take to ensure fit-for-purpose statistics, and to learn from past situations (both good and bad), to help minimise the risk of similar issues recurring. Those in KAI should consider how they can do this and whether there are existing meetings or forums that could be used to facilitate it. The first example below outlines how one team in KAI has brought together multiple areas in HMRC to learn from a past issue and to plan how to avoid similar issues going forward.

Good practice in KAI: leadership and learning from errors – Marriage Allowance

In 2019, an issue with statistics about the take up of a marriage allowance (MA) was discovered by policy stakeholders within HMRC. Between 2016 and 2019, widely used MA figures, produced by the Personal Taxes (PT) area of KAI, were not representative of what they claimed to be. In response to the discovery of this error, PT led an After-Action Review to explore what had happened, bringing together representatives from analytical, policy and operational areas of HMRC to collectively understand why, for a long period of time, HMRC misinterpreted the MA figures. The review identified several actions that are being implemented, aiming to minimise the risk of a similar issue occurring in the future. These include clarifying the sign off process for the next publication of the statistics and new audit exercises on two other statistics, which focus on understanding data flows and documentation. By bringing together multiple areas of HMRC to talk about and learn from errors, KAI helped to establish shared responsibility for the quality of statistics produced by HMRC and to improve quality management.

4.1.9. We recommend that KAI develops a network, with representatives from teams across KAI, for talking about quality in statistics and analysis. Representatives should discuss issues that different teams in KAI have faced and how they have tackled them, drawing on the Deputy HoP's issues log, as well as discussing good practice more generally. This network could be supported by the HMRC quality champion who is part of the cross-GSS quality

network and who could, therefore, relay information from other departments into HMRC and from HMRC to other departments. HMRC could look to the Welsh Government, which has a Quality Committee that supports the strategic management of quality across its statistical services: an analogous network could be effective within KAI. To persist and have an impact, this network will need to be supported by managers and senior leaders in KAI, who make it clear that it can and should be prioritised alongside team-based work. The following case study demonstrates how an informal network can be self-sustaining and bring about positive change.

Good practice in HMRC: knowledge sharing

An informal data science network exists across HMRC, supported by the HMRC Data Exploitation team. This provides a platform for hosting discussions and sharing information about events and resources. It allows staff to troubleshoot problems and ask for advice on software, statistical methods and career development. The network has resulted in self-organised groups of volunteers who have worked on HMRC-wide developments to standardise practices – for example, the recent development of an R package to automatically create charts that follow the HMRC style requirements.

4.1.10. The Lead Statisticians could also offer one another more support on issues related to quality and take a greater, more formal role in communications about quality between areas. At present, the Lead Statistician role is not consistently defined across KAI, and those fulfilling it differ in the extent to which they are involved with the quality management and assurance of official statistics in their area. It could be beneficial to define the expectations of the role more clearly. Again, if Lead Statisticians are to do more, then the Lead Statistician role needs to be recognised as an important and valued part of an individual's job, to which they can dedicate specified resource.

4.2. Working with data providers and quality assurance of input data

Statistics producers should establish and maintain constructive relationships with those involved in the collection, recording, supply, linking and quality assurance of data. Effective working relationships between parties help ensure that statistics producers understand the quality of the data used to produce their statistics and are better equipped to identify, anticipate and avoid problems that can arise from data inputs.

Findings on working with data providers and quality assurance of input data

4.2.1. Most of the official statistics produced by KAI are based on administrative data that come from HMRC's own data systems, which were not designed with statistical analysis in mind. We also note that, like many other large government departments, most of HMRC's underlying data and IT systems have been in place for many years, and do not have the same functionality or flexibility that we see in newer IT systems. This means that taking extracts of data for statistical and analysis purposes (and linking data between systems) can be difficult. As a result, there can be a significant overhead in terms of resource input for the KAI analysts to ensure that they have access to the right data extracts from administrative systems, in the correct format for their analytical work, including the production of official statistics.

4.2.2. During our discussions with producers in KAI, we heard many times that the biggest risk to the quality of statistics comes at the beginning of the production process, when data for official statistics are extracted from HMRC IT systems, transformed so that they can be analysed in KAI's statistical analysis programs and then loaded into these programs (the Extract Transform and Load (ETL) stage). Those working in KAI do not always have the relationships they need with other parts of HMRC who supply them with data to enable them to fully understand the nature of the data they are working with. Simply put, this was the origin of the issue with HMRC published statistics on Corporation Tax, which triggered this review.

4.2.3. Our review found that sometimes producers are not curious enough and do not ask enough questions about data they receive from other parts of HMRC. They do not always consider how the data have been collected, the strengths and limitations of the data, the quality checks that have been carried out by the suppliers, nor the controls and sign-off processes that are in place around the data. Those in KAI need to ask these questions (and ensure that they have built relationships that enable them to ask challenging questions) to understand the quality of the input data they are using, whether they arise from within HMRC's own systems or elsewhere.

4.2.4. Some analysts working in KAI feel that official statistics are not always considered a priority by other areas of HMRC who provide data to KAI, although this is improving after recent, public issues with official statistics. Building stronger relationships with data providers offers the opportunity to improve understanding about the work KAI does and to elevate the status of official statistics in HMRC. The HoP and the Director of KAI represent KAI on the main data governance boards within HMRC, which helps to make sure that KAI's data needs are considered when changes to data systems are made. It is important that those representing KAI and official statistics are included in discussions at all levels that relate to data, or to the IT systems that support data, across HMRC. This will help mitigate ongoing challenges that include maintaining access to data that feeds into official statistics and being informed of changes that impact the flow, content or format of these data. Not having a voice in discussions, or being-overshadowed by operational or compliance users, would be a risk to KAI being able to continue to produce official statistics that are essential to people within and outside of HMRC. Having a voice in discussions and decisions relating to data and IT systems will continue to be really important as HMRC modernises its data systems.

Recommendations on working with data providers and quality assurance of input data

Recommendation 3: In some areas, there is the need for closer working with teams providing data to KAI. The responsibility for establishing and maintaining effective relationships should be shared between KAI and its partners.

4.2.5. KAI's work is critical for other areas in HMRC, and for users of its statistics outside of the organisation: issues with KAI's statistics can impact operational delivery in HMRC and damage the reputation of HMRC as a whole. Therefore, the responsibility for establishing and maintaining effective working relationships must be shared between KAI and the other areas of HMRC that it receives data from. In some cases, the delivery of data may benefit from being supported by formal documentation (such as Memoranda of Understanding) and sign-off processes. The following example demonstrates improvements made in HMRC since the Corporation Tax error.

Good practice in KAI: building stronger relationships with data providers – Corporation Tax

Before the significant Corporation Tax receipt statistics error was discovered, colleagues in KAI and Corporate Finance did not have an ongoing dialogue about the data. Since the error was discovered, the KAI and Corporate Finance teams have been working closely together to ensure greater understanding of data flows between them. KAI has agreed that Corporate Finance will share process maps that show how the data are collected and processed before they reach KAI, and the two areas are currently agreeing a formal sign off process in Corporate Finance for data that are provided to KAI. These actions mean analysts in KAI will have a much better understanding of the data they receive and will enable more-effective ways of raising and resolving queries. Closer working between the two areas of HMRC has also raised the profile and importance of having effective processes in place to manage the production of statistics and ensure that effective end-to-end quality assurance is undertaken.

4.2.6. It is vital that those producing statistics in KAI make sure they understand the nature and quality of input data they are receiving for use in their statistics and that they plan and carry out appropriate quality assurance checks before using them. Those producing official statistics using data from operational parts of HMRC could consider if they can sit with operational colleagues on occasion, to see the systems and data 'on the ground', to better understand the data they receive. Those providing data to KAI must provide documentation about the data (such as metadata and data dictionaries). This will help KAI analysts understand what the data are and what quality checks have been completed on them.

4.2.7. We recognise that some of the statistics produced within KAI have many data sources and it is not always easy to build relationships with all those who are involved in the process of producing and transferring data. In addition, the IT systems and 'data journeys' may be complex and challenging to map and understand. However, HMRC needs to rise to this challenge, and some areas of KAI are starting to work very well with data providers, such as HMRC Corporate Finance. There will be opportunities for knowledge sharing from these areas to others where relationships are less developed: the following example shows where an area in KAI has established effective working relationships with data providers.

Good practice in KAI: working effectively with data providers – Child Benefit Statistics

The Statistics, Data and Policy team, in the Benefits and Credits area of KAI, produces annual National and official statistics about people claiming Child Benefit. Child Benefit is undergoing significant transformation as the benefit will be moved onto a new operational platform, with HMRC taking ownership of all Child Benefit data from the Department for Work and Pensions (DWP). The KAI team has established an effective working relationship with both the HMRC team leading the transformation and DWP. The KAI team is in weekly contact with the HMRC team and DWP team via meetings and email to ensure both HMRC's and DWP's analytical requirements will be met. These discussions reduce the risk of issues with the data used to produce the statistics because any changes to, or potential issues with, the incoming data are flagged to KAI and DWP users early on. This means the Statistics, Data and Policy team is prepared for changes to the content or structure of the data and has had time to look into and agree how to handle any potential issues in advance of future statistical releases.

4.2.8. Those in KAI may find it helpful to refer to OSR Guidance on Quality Assurance of Administrative Data (QAAD), which provides guidance about how to assure the quality of data from administrative systems and includes a list of the kind of questions producers should be asking themselves and their data providers, as well as case examples of how other producers have applied QAAD to enhance the quality of their statistics.

4.3. Quality assurance of analysis and publications

Quality assurance (QA) is a very important part of quality management. It provides confidence to the producer, who can in turn reassure users of the statistics, that statistics are produced to a level of quality that meets intended users' needs. QA is about identifying, anticipating and avoiding the problems that can arise from data inputs or the processes used to calculate statistics in an effective and efficient manner. QA should be proportionate to the nature of the quality issues and the importance of the statistics in serving the public good, but all statistical producers need to be curious, and not take data at face value.

Findings on quality assurance of analysis and publications

4.3.1. Our discussions with producer teams revealed that the QA done on data, analysis and publications is not always adequate across teams: some teams are doing detailed, wide-ranging, documented checks that cover the whole production process, while others are not.

4.3.2. Among the teams we spoke to, the most common method for checking to see whether new statistics are plausible is to compare them to the most recent time period. Although this is an important part of QA, we would expect all teams to also be using other approaches to QA, such as: developing their understanding of the data collection and preparation process; peer reviewing and carrying out deep dives of their analyses; and being creative in identifying appropriate independent data sources that can be used to triangulate the statistics.

4.3.3. When we conducted this review, existing KAI-wide guidance on the sign-off procedure for official statistics was good, but the expectations for and guidance about QA was limited. There was very little in the QA guidance on the kinds of questions that should be asked by those producing statistics and the kinds of checks that could be carried out throughout the production process. New guidance was being developed and was launched in early 2020.

4.3.4. Programming languages like R or Python, which could improve quality management are used in only a few teams across KAI. Senior managers in KAI see the value of applying data science techniques and there have been two test cases where data science techniques have been applied to the production of official statistics, but overall the roll out of these techniques to other KAI official statistics has been rather slow.

Recommendations on quality assurance of analysis and publications

Recommendation 4: Each team should produce and maintain a process map, which shows how its statistics are produced, from obtaining the data to publication. This map should highlight risk points and mitigating actions taken to ensure the quality of the statistics.

4.3.5. For effective QA, a producer needs to understand the data journey from the initial input to a statistical publication. This can be challenging for large departments like HMRC, given the complexity and variety of administrative data systems, and emphasises the need for analysts in KAI to develop relationships with data providers. Therefore, each team producing official statistics within KAI should produce (and plan how they will maintain) a process map, which shows the steps in the production of its official statistics, from data collection to publication. Teams should use their understanding of processes to discuss and illustrate risk points throughout the process, and the mitigating actions that are taken to ensure the quality of the statistics at these points. Teams could complete this exercise alongside one another, to stimulate thinking and to uncover areas for shared learning. Annex 5: Questions to guide thinking about quality has a list of questions producers could use to guide thinking about quality at each stage of production.

Recommendation 5: There should be a more consistent approach to, and clearer expectations for, QA across KAI. QA should be documented in each production round and available for audit.

4.3.6. Clearer expectations for QA across KAI will help teams when they are planning appropriate QA for their statistics. New, KAI-wide QA guidance has recently been developed and was launched in early 2020, which we think will be really beneficial. Making sure people in KAI are aware of this guidance and bought into it will be vital for ensuring its effectiveness. The Code of Practice for Statistics articulates a range of aspects of QA in the Quality pillar that producers should refer to when planning QA. The GSS 'Quality statistics in government' strategy, produced by the GSS Best Practice and Impact Division, also supports statistical producers in meeting the quality requirements of the Code and provides helpful steps to consider in QA. QA carried out in each production round should be documented and available for audit.

4.3.7. In each production cycle, all teams should be considering and implementing appropriate checks that help them to judge the accuracy and reliability of input, processed and output data. These could include checks on item completeness (or 'missingness'), record completeness, data coverage, precision/uncertainty and sensitivity/specificity. Teams should also consider whether they can compare their statistics to independent data sources (within or outside HMRC) to gauge if their numbers look plausible. We also recommend that all teams complete 'wash ups' after publications, to consider what went well and what could be done better; perform 'deep dive' reviews of analysis and methods (including QA methods) on a systematic basis; and that a programme of peer review of analysis is implemented across teams or areas. The following two examples explain the steps some teams in KAI have taken to effectively assure the quality of their statistics.

Good practice in KAI: reviewing legacy code – Creative Industries Statistics

The team producing the annual official Creative Industries Statistics decided to review legacy SAS code that had become complex and difficult to check. This led to the team identifying figures that were being excluded by the code along with problems with the way projects (e.g. films) were counted when there were multiple projects on one tax relief claim. The analysts talked to the HMRC Creative Industries unit which supplies the source data to ensure a correct understanding of how the claims are recorded. The team then re-wrote the SAS code, testing it thoroughly and creating a flowchart to illustrate the process. The revised SAS code was used to produce the 2019 release of the Creative Industries Statistics, improving the quality of the statistics. The simpler code is now easier to check and maintain, reducing the risk of future issues.

Good practice in KAI: robust QA processes – Child Benefit Statistics and Income tax deducted from pay statistics

Two of the teams we spoke to in KAI have particularly thorough and robust quality assurance processes: the Child Benefits team and the External Data team, both in the Benefits and Credits area of KAI. Each team has developed a bespoke checklist which details the checks that are carried out when the publication is produced. This provides assurance for themselves and their managers when it comes to publication sign-off and is a useful tool for new starters to guide them through the process. The checks that the teams carry out include comparisons to alternative data sources – for example, when constructing statistics on the number of migrants claiming HMRC benefits broken down by country, the responsible team compares trends with estimates of the migrant population published by the ONS, to check for consistency – automatically flagging changes over time that exceed a pre-specified threshold, a second check of code by another analyst and using their expert knowledge of the data and the policy context to investigate and understand trends.

Both teams have carefully considered their overall production processes and identified points that pose the highest risk of errors, allowing them to focus checking or improvement efforts. For example, in response to an error that was spotted by an external customer in the Child Benefits publication, the team moved final calculations from Excel into SAS. This has minimised manual steps in the production process and therefore reduced the overall risk of error.

Recommendation 6: Drawing on its relationship with the HMRC Data Exploitation team, KAI should move towards greater use of data science methods and tools.

4.3.8. Programming languages like R or Python should be employed where it is judged they could enhance the quality of statistics; for example, by helping teams to more efficiently and effectively perform checks on input data and to remove manual steps from the production process. Version control software would help teams manage multiple versions of code and provide a guaranteed audit trail of changes to code over time. Using these tools would also allow easier peer review of analysis.

4.3.9. Two teams in KAI have already been used as test cases to implement these practices by developing 'Reproducible Analytical Pipelines'[1]. These teams, and other departments across the GSS, have demonstrated benefits such as effectively managing versions of code, removing error-prone steps in the production process and creating efficiencies, as well as the ability to produce more dynamic final publications, which the user can explore. It should be considered how the learning from these cases can be spread effectively and sustainably across teams and areas in KAI: to support this, there is the opportunity to work with the HMRC Data Exploitation team, which should be explored. This is a specific example of where support from senior leaders will be important to ensure that a programme of work to apply RAP to suitable publications is developed, as well as to make it clear to those producing official statistics that this work can be prioritised as part of day-to-day statistics production. The example below illustrates the work that is already ongoing in KAI.

Good practice in KAI: use of data science methods and tools

As an initial step in moving selected official statistics publications into fully reproducible pipelines, two teams have transferred the graphics design and production elements of the production process into R and published these outputs. Statistical bulletins for the annual Corporation tax National Statistics and Measuring tax gaps official statistics have been produced in this way. The teams collaborated with the wider data science community to proactively develop a bespoke R package: the 'HMRCstyles' package, which improves the quality assurance process by reducing the amount of code needed to produce figures that are both fully accessible (i.e. tested for colour vision deficiency) and that adhere to HMRC Corporate Communications visual design guidelines. Use of the HMRCstyles package resulted in improved quality assurance, as well as more consistent graphic design, across the publications. A number of other teams across KAI are in the process of applying this method to their own publications.

[1] For those wishing to learn more about Reproducible Analytical Pipelines (RAP), there is a good explanation in the Quality statistics in government strategy. This document clarifies what 'Reproducible Analysis Pipelines' are, the benefits of RAP and provides information on where producers wishing to learn more about implementing RAP can get further support.

4.4. Published materials relating to quality of HMRC official statistics

To make sure people using government statistics are well informed about the quality of published statistics and how they can be used, and to avoid potential misuse, producers should always explain how they assure themselves that statistics and data are accurate, reliable, coherent and timely. As part of this, the strengths and limitations of the statistics and data should be considered in relation to different uses and explained alongside the statistics. Any scheduled revisions or unscheduled corrections that result from errors should be explained alongside the statistics, being clear on the scale, nature, cause and impact.

Findings on published materials relating to quality of HMRC official statistics

4.4.1. While some HMRC statistical publications have up-to-date information on quality management and quality assurance within them, much of the easily found published information on GOV.UK about methodology, quality and administrative data sources relating to HMRC official statistics does not fully explain to users the strengths and limitations of the statistics and data, or how quality is managed across the whole production process. We note that many of HMRC's published documents about quality have not been updated since 2013. Being transparent about limitations, as well as providing helpful support to users, would enhance the trustworthiness of HMRC as a statistics producer.

4.4.2. HMRC publicly announces delays to its official statistics on GOV.UK. From the homepage for HMRC statistics, these are accessed through a document titled HMRC announcements, which is found in a link titled Announcements: changes to the publication schedule, which is itself found under the heading Schedule of Updates. It is not clear from the main landing page that changes to the publication timetable, nor the reasons for them (such as forthcoming corrections, as was the situation with Corporation Tax statistics in September 2019) would be found by following that series of links, and it is not obvious that this is where users should be looking to learn about issues with HMRC's published statistics. Furthermore, the information provided within the announcements document tends to be brief. There is a risk to the perception of HMRC's trustworthiness as a producer of official statistics, that could be reduced through clearer and more prominent announcements.

Recommendations on published materials relating to quality of HMRC official statistics

Recommendation 7: HMRC should review the published information about the quality of its official statistics and update to ensure that it fully explains how HMRC producers assure themselves that their statistics are accurate, reliable, coherent and timely.

4.4.3. To improve trustworthiness and transparency to users about the quality of published official statistics, all teams should review the information they currently provide relating to the quality of individual official statistics. Published information should explain how HMRC producers assure themselves that their statistics and data are accurate, reliable, coherent and timely, and consider all stages of the production process: from data collection to publication. Producing process maps, as we recommended earlier, would help equip teams to produce this information.

4.4.4. Teams should apply the OSR Quality Assurance of Administrative Data toolkit as they seek to improve their QA processes and the information they provide about the quality of their data and statistics. The OSR website provides a large number of case examples teams can draw on, which highlight different ways that statistical producers have applied the QAAD guidance. There are also two case studies on the Code of Practice website that demonstrate how the Office for National Statistics and NHS Digital have fulfilled their commitments to publishing clear and detailed information about data quality assurance and background quality. Most of the official statistics produced by KAI are based on administrative data, and teams may also find it helpful to look at the user guide for the Legal Aid Statistics, produced by the Legal Aid Agency, which presents a thorough discussion of data sources and quality in relation to use of complex administrative systems, and a section

on revisions, which clearly explains changes to administrative systems or methodology. The Welsh Government also provides very good information on the quality of its Homelessness statistics, which are based on administrative data, in a user guide.

Recommendation 8: HMRC should publish information in a transparent way that is easy to find.

4.4.5. The content of notices about issues or delays to published statistics should be improved, to ensure that users always have information that enables them to understand why an issue has arisen and the expected timeframe for resolution. The visibility of published notices should also be reviewed: notices could be published on the GOV.UK pages of the specific statistics to which they relate, for example, as well as in the Schedule of Updates.

4.5. Analytical resources and training

To produce high quality statistics, and to be able to drive improvements and innovation in statistics, people producing them need to be appropriately skilled, trained and supported in their roles and professional development. There also needs to be sufficient human, financial and technological resources to deliver statistical services that serve the public good.

Findings on analytical resources and training

4.5.1. Senior managers in the four areas of KAI we focused on felt that they have enough resource to do the core work that they are currently required to do. However, from talking with senior managers and more junior producers, we get the sense that some analysts feel that they do not have the space to invest in developing and improving their statistics. This includes building and maintaining relationships with data providers and doing wider research to help them fully understand their data and statistics, and the wider context around them; undertaking and documenting the highest levels of QA; and introducing new tools. We found that one area of KAI – Indirect Taxes Customs & Co-ordination – had freed up resources by reviewing and updating its portfolio of publications, but this activity has not been undertaken systematically across KAI.

4.5.2. We heard that relatively frequent rotation of staff, mainly at more junior levels within KAI, can be a barrier to team members having a good understanding of the wider context around the statistics they work on. It can also mean that teams and managers have to invest a lot of time in recruitment and getting new recruits up to speed. There is a trade-off here for HMRC to balance as, on the other hand, allowing staff to rotate within KAI means they are more likely to stay in HMRC, rather than moving to other Government Departments. We found that one area's approach was to have a small, flexible team without fixed outputs that could be deployed as priorities dictate.

4.5.3. With respect to skills and training, the teams we spoke to felt they have good access to training within HMRC, including statistical packages such as SAS (which is its default statistical package) and R.

4.5.4. While there is great enthusiasm among statistical producers to learn new programming languages and apply the principles of Reproducible Analytical Pipelines, there are challenges to be overcome in implementing them. These include lack of time to learn and implement the new techniques (even if training is available), along with the fear, more among senior managers, that knowledge management won't be good enough to maintain developments in the light of staff turnover. We also heard that innovative or development work is often given to the newest and most transient members of a team, which, without the right staff resilience in place, can mean developments are unlikely to endure.

Recommendations on analytical resources and training

4.5.5. To improve quality management in KAI there is a wide range of initiatives that will need further development, embedding and managing. HMRC should consider the changes that need to be implemented and come to its own conclusions about whether additional resource will be needed in KAI or the areas of HMRC who work with KAI, or if the changes can be achieved by using existing resources differently. In either case, there may be an up-front requirement for additional resources to review and change current practices.

Recommendation 9: Each area of KAI should review its current publications, with a view to reducing the number or size of publications, to ensure analytical resource is being used effectively.

4.5.6. As part of assessing the quantity and allocation of currently available resource in KAI, we recommend that all four areas of KAI should review, in consultation with users, their current publications. Teams within KAI may be able to reduce their total number or size of publications; by reducing the number of times a bulletin is published during a year, for

example. This could free up resource that could be invested in activities that support the quality and development of remaining publications and, at the same time, help HMRC fulfil its commitment to producing statistics that are of public value and meet the needs of users. One area of KAI has previously performed this kind of review, to great effect – more detail is given in the first example below. Another area of KAI is using web-page analytics and public consultations to better understand which of their statistical outputs are the most read and how their statistics are used, in order to refine them and make them more valuable to users – more information is given in the second example below. These approaches could be used by other areas of KAI to support decisions in a review of publication outputs. Clear messages of support from senior managers about 'doing less better' could help ensure review exercises are undertaken and effective.

Good practice in KAI: reviewing publications – Indirect Taxes Customs & Co-ordination

In the Indirect Taxes Customs & Co-ordination (ITCC) area of KAI, the Deputy Director, the Lead Statistician and the head of the team producing the majority of the area's publications conducted a review of their official and National statistics publications. They mapped out everything the team published and the frequency of those publications. After this review, an alternative programme of statistics publications was drawn up and a user consultation undertaken. The team was then able to reduce the frequency of publications (e.g. moving from monthly to quarterly releases, or quarterly to annual), which resulted in around 100 publications a year being reduced to around 70. This has freed up time to undertake more 'value added' work – investigating the reasons for changes in the numbers and developing knowledge of the data, for example. As well as improving quality assurance processes, this has meant that the team can better spot potential issues with the data and statistics. The ITCC Lead Statistician has also driven a programme of work to improve the presentation of ITCC's statistics – using a consistent ITCC format, shorter and more focused presentation of the data and use of infographics, for example.

Good practice in KAI: considering user needs in developing outputs – Personal Taxes

In support of devolved Scottish income tax, the Personal Taxes (PT) area of KAI has recently launched a new Scottish Income Tax Outturn Statistics publication. In developing this, KAI worked very closely with key stakeholders in the Scottish Government, Scottish Fiscal Commission and the Office for Budget Responsibility to ensure the publication was focused on user needs. Alongside the first publication of the statistics, a User Needs and Feedback survey has been launched to capture views from users to help refine the publication.

In order to develop a broader understanding of how people are using a range of PT statistics, the Specialist Personal Taxes branch of PT has done a detailed analysis of web analytics, such as the number of hits on different publication pages, to work out which are the most read.

4.5.7. KAI should also consider whether the use of flexible resource teams, as explained in the example below, can be extended from Direct Business Team to other areas of KAI.

Good practice in KAI: ensuring analytical resource – Direct Business Taxes

The Direct Business Taxes (DBT) area in KAI has set up a small 'flexible resource team' of three people, situated in the research team in DBT, which does not have fixed outputs. Having this flexible team means extra resource can be put into analysis and statistics produced by DBT when necessary: the team can help undertake development work, or when issues are discovered with data or statistics. For example, this year DBT trialled a new methodology to combine Research & Development data for a statistical publication and was able to set up a parallel team to use the new source and replicate the analysis. This was then used as the base dataset for the publication.

4.5.8. While there remains relatively high turnover within KAI, teams must consider how they can best protect business continuity and ensure knowledge retainment. This could include allowing for staff change in statistical production timetables and a requirement that any individual who is leaving produces a peer-reviewed handover document before they go, which is signed off by their manager.

Annex 1: HMRC Corporation Tax Statistics

The error

In autumn 2019, HMRC identified an error in the published Corporation Tax receipt statistics. The error required HMRC to correct Corporate Tax estimates for the period April 2011 to July 2019, which it did in its 'Tax and NIC receipts: information and analysis (August 2019)' statistical publication on 24 September 2019: https://www.gov.uk/government/statistics/hmrc-tax-and-nics-receipts-for-the-uk

This was a statistical reporting error, which did not affect the amount of tax paid by companies and received by HMRC, or the figures reported in HMRC's accounts.

Corporation Tax statistics need to take account of Corporation Tax Credits, which are reliefs designed to encourage various activities. They work by either reducing a company's liability to Corporation Tax, or by making a payment to the company. To produce gross Corporation Tax receipts, the value of these credits is added to the payments received. In autumn 2019, quality assurance checks identified some double counting of credit values when using internal financial data in the statistical production process.

Impact

The scale of the revisions varies between years. The revised Corporation Tax receipts figures published are £650m lower in 2011/12, £2.3bn lower in 2017/18, and £4bn lower in 2018/19.

The revisions affected the Office for National Statistics' (ONS) estimate of Public Sector Net Borrowing, which draws on the gross Corporation Tax receipts statistics. There was no impact on Public Sector Net Debt for outturn years. The ONS figures published on 24 September 2019 took on HMRC's revised receipts series, together with other data and methodology changes. The changes were explained in the ONS publication: https://www.ons.gov.uk/economy/governmentpublicsectorandtaxes/publicsectorfinance/bulletins/publicsectorfinances/august2019

HMRC Response

In response to this error, HMRC commissioned this independent review of the quality management and quality assurance processes used to generate HMRC's published statistics. The review was led by Ed Humpherson, Director General for Regulation and head of the Office for Statistics Regulation.

HMRC has also carried out internal work, including an internal audit of the circumstances around the Corporation Tax error, to fully evaluate this issue and to inform corrective actions and improvements to mitigate the risk of a similar issue in the future.

Annex 2: HMRC as a producer

HMRC as a producer

The Knowledge, Analysis and Intelligence directorate (KAI) in HMRC is responsible for providing analysis, research and statistics to inform policy development, operational design and internal HR processes. It produces and publishes around 150 statistical releases, covering over 60 topics. The publication plan for HMRC National and official statistics is published on GOV.UK. Much of KAI's work is supporting the wider tax revenue and customs functions of HMRC, beyond official statistics production. The statistical producer teams in KAI also often partner with teams in HM Treasury and the Office for National Statistics.

KAI has just over 450 full-time equivalent (FTE) posts in eight business areas (including the Deputy Director's People and Performance team) (see organogram below). Around one-third of KAI staff are involved in producing official statistics, based in five divisions: Benefits & Credits; Customer Compliance Strategy; Direct Business Taxes; Indirect Taxes Customs & Co-ordination; and Personal Taxes. Areas providing analytical insight on customer compliance account for just under a third of KAI staff (134 FTE) (in the Customer Compliance Strategy and Compliance & Debt Operations business areas). KAI Operations, Strategy & Transformation accounts for a further 58 FTE, and the Technology Management Team, 15 FTE.

Roles in producing official statistics

Analysts below grade 7 are responsible for carrying out analysis and quality assurance for a given statistic or publication.

The grade 7 lead analyst/team leader for the statistics or publication is responsible for setting the expectation of the QA that will be done and for checking it has been completed:

- The grade 7 is responsible for working with their team to construct a fit for purpose QA checklist that is tailored to their output.
- The grade 7 then decides who should carry out the QA tasks and is responsible for ensuring those tasks are carried out correctly and reviewing the results of the QA.

The grade 6 manager (who would normally oversee multiple publications) is responsible for confirming they are happy with the QA framework proposed by the grade 7 and then for sense-checking draft publications:

- The grade 6 is responsible for ensuring the grade 7s have constructed a checklist, that it is fit for purpose, and completed.
- Building on the work of the grade 7 would usually be a "fresh pair of eyes".
- Asking the grade 7 and team probing questions about how the QA had been done and looking for a consistent message ensuring differences were explained.
- Ensuring that the grade 7 has effectively planned the process, including sufficient contingency, and will consider the wider context for supporting briefing around the statistics.

The Lead Statistician's role is not to do the work of the grade 6 manager (or the DD), but to advise producers in conjunction with the HoP where necessary about Code of Practice issues and how best to handle risks. In some circumstances the lead statistician will share the QA role with the G6 manager.

The Deputy Director (Senior Civil Servant) is responsible for the sign-off of the statistics in their area and is ultimately accountable for these statistics. The DD would be expected to have an active role in signing off any briefing associated with a statistical publication and therefore needs to understand what the release is saying and whether more senior managers in HMRC need to be aware of the statistics.

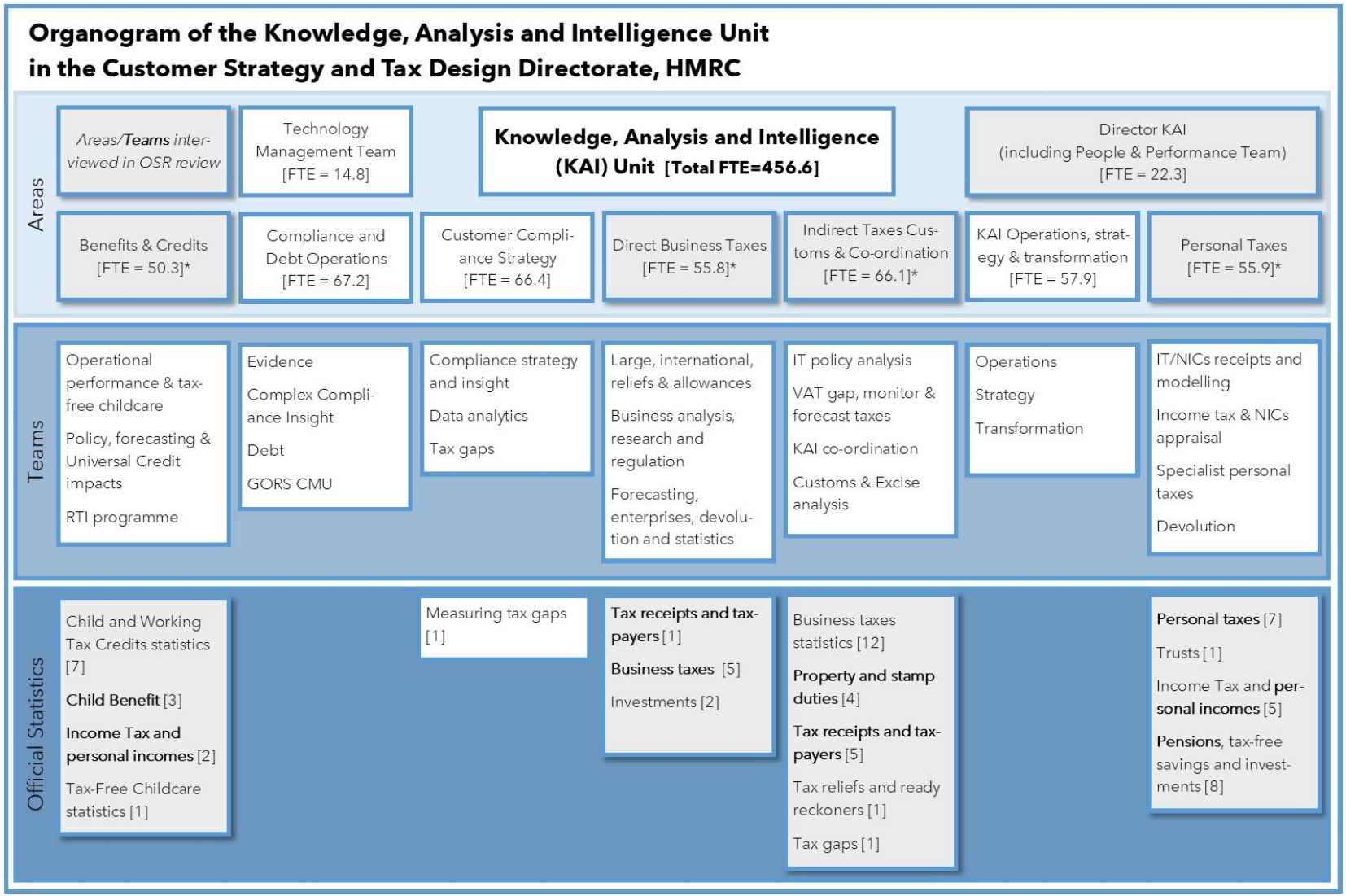
The Director of KAI is ultimately accountable for all KAI statistical publications and as a result reviews and signs off key statistical releases and associated briefing.

For publications with high public or political interest, additional levels of sign off by the Director of KAI and/or the Director General of the Customer Strategy and Tax Design Directorate are required.

The Head of Profession for Statistics advises on statistical policy and best practice, particularly where it is a first publication or where problems are being encountered.

Source: HMRC FTEs January 2020, provided by HMRC

Organogram of the Knowledge, Analysis and Intelligence Unit in the Customer Strategy and Tax Design Directorate, HMRC



Number in square brackets gives the number of sets of Official Statistics for each topic - https://www.gov.uk/government/organisations/hm-revenue-customs/about/statistics#schedule-of-updates
FTEs as at January 2020 (excluding vacancies) provided by HMRC

Annex 3: List of meetings undertaken during this review

Below is a full list of those OSR engaged with during the review. We would like to express our gratitude to everyone who gave their time to speak with us, for their cooperation and openness.

HMRC

Senior leaders and statistics producers in KAI

Director, KAI

HMRC Head of Profession for Statistics / Deputy Director, Direct Business Taxes (DBT)

Deputy Head of Profession for Statistics

Deputy Director, Personal Taxes (PT)

Deputy Director, Benefits and Credits (B&C)

Deputy Director, Indirect Taxes Customs and Coordination (ITCC)

Lead statistician, DBT

Lead statistician, PT

Lead statistician, B&C

Lead statistician, ITCC

Corporation Tax Receipts, Liabilities and Bank Levy team, DBT

Creative Industries/Research and Development Tax Credits team, DBT

PAYE and Corporate Tax Receipts from the Banking Sector team, DBT

Marriage Allowance team, PT

Personal and Stakeholder Pensions team, PT

Personal Incomes team, PT

Child Benefits team, B&C

Income Tax Deducted from Pay team, B&C

HMRC Receipts team, ITCC

Property and Stamp Duties team, ITCC

North West Statisticians group

Teams outside KAI

Assistant Director, Data Exploitation, Chief Data Office (this team is supporting KAI as it moves towards greater use of data science methods and tools in its production of official statistics)

Teams and organisations who provide data to KAI for use in official statistics

Data providers to the Corporation Tax Receipts team, Corporate Finance, HMRC

Data providers to the Creative Industries team, Customer Compliance Group, HMRC

Data providers to the Marriage Allowance team, KAI and Chief Data Office, HMRC

Data providers to the Research and Development Tax Credits team, Customer Compliance Group, HMRC and Capgemini

External Stakeholders

HM Treasury: Head of Profession for Statistics

National Audit Office: Director, Financial Audit; Audit Manager, Financial Audit

Office for Budget Responsibility: Deputy Chief of Staff; Team Leader; Senior Analyst, Fiscal Team

Office for National Statistics (ONS): Public Sector Finance Statistics production team

Office for National Statistics (ONS): Quality Centre, Best Practice and Impact Division

Scottish Fiscal Commission: Head of economy and tax forecasting; Senior Analyst; Senior Analyst

Annex 4: Key findings and recommendations of the review

The table below summarises the key findings and recommendations of this review of the principles and processes underpinning the quality of HMRC's published official statistics. More detail about the findings and recommendations can be found in section 2 and section 4 of the main report.

Table 1: Key findings and recommendations

	Finding	Recommendations
Leadership and oversight of quality management in KAI		
1	Senior leaders in KAI do not always give sufficient priority and resource to improving quality management and quality assurance. (Finding 4.1.4)	Senior leaders in KAI should champion and support developments and innovations that can enhance the quality of official statistics. (Recommendation 4.1.7)
2	There is little opportunity for routine sharing of good practice with respect to quality, or learning from past issues in statistics, across KAI. (Finding 4.1.6)	Knowledge sharing should be prioritised at all levels of KAI, and more-effective mechanisms built to support this. As part of this, a cross-KAI quality network could be put in place, with senior support. (Recommendations 4.1.8 to 4.1.9)
Working with data providers and QA of input data		
3	Producers in KAI do not always have effective relationships with data suppliers, many of whom are from other parts of HMRC, and do not always fully understand the nature and quality of the data they receive. (Finding 4.2.1)	In some areas, there is the need for closer working with teams providing data to KAI. The responsibility for establishing and maintaining effective relationships should be shared between KAI and its partners. (Recommendations 4.2.5 to 4.2.6)
QA of analysis and publications		
4	Not all teams consider the whole data journey, from the initial data input to a statistical publication when carrying out QA. (Finding 4.3.1)	Each team should produce and maintain a process map, which shows how its statistics are produced, from data collection to publication. This map should highlight risk points and mitigating actions taken to ensure the quality of the statistics. (Recommendation 4.3.5)
5	The application and completeness of QA done during the statistical production process is not always adequate across teams. (Finding 4.3.1 to 4.3.3)	There should be a more consistent approach to, and clearer expectations for, QA across KAI. QA should be documented in each production round and available for audit. (Recommendations 4.3.6 and 4.3.7)
6	More teams could be using data science methods and tools to effectively manage and peer review code, remove error-prone steps in production and create efficiencies. (Finding 4.3.4)	Drawing on its relationship with the HMRC Data Exploitation team, KAI should move towards greater use of data science methods and tools. (Recommendations 4.3.8 and 4.3.9)
Published materials relating to quality of HMRC official statistics		
7	Much of the published information on quality of HMRC official statistics does not fully explain to users the strengths and limitations of the statistics and data, or the production process. In some cases, it does not appear to have been updated since 2013. (Finding 4.4.1)	HMRC should review the published information about the quality of its official statistics and update to ensure that it fully explains how HMRC producers assure themselves that their statistics are accurate, reliable, coherent and timely. (Recommendations 4.4.3 and 4.4.4)
8	Published information about issues with and delays to official statistics can lack detail and is not easy to find. (Finding 4.4.2)	HMRC should publish information in a transparent way that is easy to find. (Recommendation 4.4.5)
Analytical resources and training		
9	Analysts producing statistics in KAI can struggle to find time to invest in developing and improving their statistics. Analytical resource that could be used for this may be being spent on producing publications with more content, or that are produced more frequently, than needed by users. (Finding 4.5.1)	Each area of KAI should review its current publications, with a view to reducing the number or size of publications, to ensure analytical resource is being used effectively. (Recommendation 4.5.6)

Annex 5: Questions to guide thinking about quality

This annex includes questions that analysts producing statistics in HMRC, and other organisations, can use to guide their thinking about quality at each stage of production. This list, which is adapted from the series of questions we asked teams in HMRC who produce statistics during this review, is not a checklist, and is designed to be used alongside guidance within producer organisations, and alongside external resources such as the Code of Practice for Statistics and the Quality Assurance of Administrative Data Toolkit, which are listed at the end of the Annex.

Statistics should be produced to a level of quality that meets users' needs, and quality assurance (QA) should be proportionate to the nature of the quality issues and the importance of the statistics in serving the public good.

Understanding the production process

What are the steps in your statistical production process, from acquiring the data to final publication? Can you map out the "data journey"? Why is it done in this way?

Where are the highest risk points for errors in the process? What measures do you or could you take to mitigate risk at these points?

How much time does your publication take to produce and how is this time split between data collection, analysis, report preparation and QA? Does the current balance feel effective?

How do you know that the publication is ready? Who is responsible for final sign off?

Tools used during the production process

What analytical tools do you use during the production process? Are they the best for the job?

How many manual steps are there in the process (e.g. updating cells in spreadsheets, moving data between software or copy-paste steps)? Could these be reduced to minimise the risk of error?

Are any parts of the publication process automated? If so, how do you ensure that these are correct and can be inspected and understood by other staff or new members in the team?

Receiving and understanding input data

When and how do you communicate with your data provider(s)?

Does your data provider have a good understanding of how and why you are using their data?

Is there a formal agreement in place that specifies when, what and how the data will be received? If not, do you think this would be helpful?

Do you know what quality checks are carried out on the data before you receive them?

How do you work with your data provider when your data requirements change?

How do you know if your data provider makes a change to their systems or processes, which could impact the data you receive and/or the statistics you produce?

What are the strengths and limitations of the data used in your publication? Are these communicated to people using your statistics?

Quality assurance

What do you feel is done well with regards to QA in your team? What could be better?

How do you ensure that input data are correct and in the expected structure and format?

How do you assure yourselves that analysis carried out is correct?

If you find anomalies or unusual trends in the data, what steps are taken to investigate them?

Is your code or analysis ever peer reviewed by someone outside your team or organisation?

Version control and documentation

How do you ensure that analysis is auditable and can be inspected and understood by colleagues?

Could you reproduce the analysis and output from a previous publication?

If changes need to be made to any code or analysis, how are these documented? Are changes checked by another member of the team?

Issues with the statistics

What happens if you find a mistake in the data/your publication? How is it rectified?

What steps would you take to minimise the chance of a similar error happening again?

Other resources available

The Code of Practice for Statistics The Office for Statistics Regulation Code of Practice for Statistics provides producers of official statistics with the detailed practices they must commit to when producing and releasing official statistics. The Code articulates a range of aspects of QA in the Quality pillar that producers should refer to when planning QA.

Quality Assurance of Administrative Data (QAAD) Toolkit (February 2019) This toolkit from the Office for Statistics Regulation is intended to help statistical assessors review the areas of practice for the quality assurance arrangements of administrative data used to produce official statistics

QAAD Questions – what do I need to ask? Use this guide from the Office for Statistics Regulation as a prompt to help you find out more about your administrative data sources and any associated quality issues.

The GSS 'Quality statistics in government' strategy, produced by the GSS Best Practice and Impact Division, also supports statistical producers in meeting the quality requirements of the Code and provides helpful steps and techniques to consider in QA. This guide includes an example of a "data journey".