Data Quality

City of York Council

Internal Audit Report 2014/15

Business Unit: Office of the Chief Executive, Business Intelligence
Responsible Officer: Programme Director, Business Consolidation
Service Manager: Group Manager, Shared Intelligence Bureau
Date Issued: 11 June 2015
Status: Final
Reference: 11040/006

	P1	P2	P3
Actions	**0**	**5**	**0**
Overall Audit Opinion	Reasonable Assurance		

Summary and Overall Conclusions

Introduction

Public services need reliable, accurate and timely information with which to manage services, make decisions and account for performance.

Performance Indicators provide key information to enable service performance to be monitored. They are an important driver of improvement and allow key stakeholders to judge how well a service is performing.

It is essential for the users of performance information to be able to place reliance on the figures supplied and good quality data is the essential ingredient for reliable performance and financial information. However, a balance must be achieved between the need for information and the cost of collecting the supporting data with the necessary accuracy, detail and timeliness.

The council's performance indicators cover a wide range of services and are reported to the Corporate Management Team, Members and the public on a quarterly basis on scorecards for each of the council's corporate priorities.

15 council indicators have been selected for detailed review, covering all directorates. These were selected from a risk assessment taking into account a range of factors, including: where method of calculating the indicators has changed; significant variances have occurred from one period to another; known issues with the collection or reporting of the indicator; the issues around the indicators are of high political or strategic importance.

Objectives and Scope of the Audit

The purpose of this audit was to provide assurance to management that procedures and controls within the system will ensure that for each indicator:

- the data set needed to calculate was complete, accurate and relevant for the calculation of the indicator;
- the data was correctly processed in order to calculate the indicator;
- the department reviews the data gathering process and final output figures to confirm they are accurate.

The audit also considered whether appropriate corporate management arrangements for data quality are in place and being applied in practice. It did not consider the wider performance implications around the PIs chosen, or the extent to which the chosen PIs meet the planned aims and objectives of the council.

Key Findings

The table at Annex 1 summarises the findings for each performance indicator tested. No significant errors were found in the accuracy of the figures reported for each performance indicator. Where there were differences between the figures originally reported and the audit testing there were satisfactory explanations for this (most commonly timing differences).

Although no significant errors were found in the figures reported there were a number of issues noted that increase the risk of errors being made and thus limit the level of assurance that can be given:

- There was a lack of consistency in the controls in place to check the accuracy of data produced and a lack of corporate guidance available to officers on what controls should be in place to ensure data quality.
- The production of the performance indicator is heavily reliant on the knowledge and experience of the officer producing them, who, in many cases, have been producing the indicator for many years.
- There were no formal definitions of the indicator or procedure notes / guidance on the data collection and calculation of the indicator.
- There are significant manual processes involved in transposing the data into spreadsheets that calculate the performance indicator and are used to produce the management reports.
- There was no review of the performance indicator against the base data so this process is reliant on individuals so errors may be more likely to go undetected.

Individually these issues do not seem to have resulted in incorrect performance being reported. However, when taken in combination they are likely to increase the risk of errors and this will also be significantly increased in the event that key staff leave the council.

Overall Conclusions

It was found that, while the individual indicators reviewed had been produced accurately, there is currently a lack of corporate arrangements in place in relation to data quality and this may leave the council exposed to an increased level of risk. An acceptable control environment is in operation but there are a number of improvements that could be made. Our overall opinion of the controls within the system at the time of the audit was that they provided Reasonable Assurance.

1 Data Quality policy and procedures

Issue/Control Weakness	Risk
Lack of corporate data quality policy and procedure.	Without a corporate data quality policy there is an increased risk that standards will not be known and adherred to by those responsible for processing data and producing performance indicator figures. Ultimately, if data quality is not maintained then performance may be incorrectly reported and decisions could be based on incorrect information.

Findings

There is no corporate Data Quality policy or procedure that sets out the principles, standards and responsibilities for the production of Key Performance Indicators (KPIs) (and data quality more generally). The most recently available policy is dated 2009 and whilst this contains useful and relevant information on data quality principles it is written in the context of the old Comprehensive Area Assessment (CAA) process and reporting of PIs to the Audit Commission. This policy is not available on the intranet and officers responsible for producing PIs do not make any reference to it. This means that there are no formal corporate standards for ensuring data quality.

Agreed Action 1.1

The business intelligence (BI) hub has now brought the production of all performance indicators into one central team. The structure itself enables more supervision and checking of indicators and all KPIs are signed off by managers.

Training needs for all members of the BI hub have been identified through the performance and development review (PDR) process.

Priority	2
Responsible Officer	Group Manager - Shared Intelligence Bureau
Timescale	Action completed

2 Producing and Reviewing Performance Indicators

Issue/Control Weakness	Risk
Lack of management controls in production of individual indicators.	A lack of controls in the production of performance indicators increase the risk of errors being made and not being identified. Ultimately this may lead to performance information being reported inaccurately and decisions being made on the basis of incorrect information.

Findings

There was a lack of consistency in the controls in place to direct and check the accuracy of data produced. For many indicators tested the production of the performance indicator is heavily reliant on the knowledge and experience of the officer producing them, who have often been producing the indicator for many years (this was noted to be particularly apparent in 4 of the 12 areas tested). Additionally, in half of the areas tested there were no procedure notes, guidance or formal definition of the indicator. Processes involved in collation of the indicator often included manual calculations or transposition of figures to spreadsheets. Taken in combination these factors mean that in the event that key officers leave the council the risks to data quality would be greatly increased.

In addition, in 5 of the 12 areas tested the performance indicators were produced entirely by one person and only reviewed when they were reported to managers. Whilst this might lead to figures being queried and re-checked this does not happen in a systematic way and relies on the reported figure being significantly different to what expected performance would be. There was no review of the performance indicator against the base data to check its accuracy before being reported to managers.

Agreed Action 2.1

As noted in Action 1.1, the new structure enables closer supervision and checking of indicators and more officers are involved in the production of indicators, so the risks associated with individual officers being solely responsible for the production of indicators and the lack of management checks on indicators is reduced.

Priority	2
Responsible Officer	Group Manager - Shared Intelligence Bureau
Timescale	Action completed

Agreed Action 2.2

The collation of indicators is now automated through a single data table maintained by the BI hub (the 'KPI machine'), which reduces the risks from manual calculations and manual transposition of figures into different locations.

The 'KPI machine' also runs a number of automated reports to identify: unexpected variances; direction of travel anomalies; missing data. These are reviewed and investigated before performance indicators are signed off.

Priority	2
Responsible Officer	Group Manager - Shared Intelligence Bureau
Timescale	Action completed

Agreed Action 2.3

All existing procedure notes have been brought into one central location where they are available to all members of the BI hub.

All existing procedure notes will be reviewed, and updated where necessary. Procedure notes will be written for any performance indicators where they do not already exist.

Priority	2
Responsible Officer	Group Manager - Shared Intelligence Bureau
Timescale	September 2015

Agreed Action 2.4

The longer term strategy for the development and implementation of new IT systems is for data quality controls to be built into systems and for information to be able to be extracted directly from systems with minimal manual intervention.

This strategy has been agreed between the Group Manager (Shared intelligence bureau) and the ICT Business Engagement Manager. It will be implemented on a system by system basis.

Priority	2
Responsible Officer	Group Manager - Shared Intelligence Bureau
Timescale	Immediate and ongoing

Annex 1

Indicator Reviewed	Summary	Weaknesses / Issues requiring attention
Adults with learning disabilities: a) in settled accommodation b) in employment	Source data is entered by social workers following annual reviews and reports are run identifying all clients in employment or settled accommodation. The results of these reports are then manually entered into a spreadsheet to calculate the indicator. One minor error was found, which was likely to have been the result of an error in the manual input of figures to the final management report. This was not material and overall the indicator had been reported accurately from the source data. A new officer now collates this data and reported that the lack of any procedure notes made collection of the indicator difficult for the first few months.	There is no explicit management check of the PI figures. There are no procedure notes detailing how to collect the data and calculate the indicator.
Number of Common Assessment Frameworks (CAF) initiated	The initiation of CAFs is reported to the Children's Advice Team and recorded on the E-trak database. Termly interventions are conducted with practitioners based in schools to try and ensure all CAFs have been reported and data cleansing exercises are undertaken annually. The indicator is calculated from the records on the E-trak system and had been reported accurately for the period tested.	There is no second officer review of the figures but the interventions and data cleanses mitigate are compensating controls. The service is currently reviewing all programmes to ensure they are CAF compliant.
Average number of weekly CYC acute delayed discharge	Data is collected on an NHS database and processed by NHS England. A council representative attends meetings with the NHS representatives to agree which delayed discharges from hospital are mainly the responsibility of the council. There are detailed procedure notes to ensure data is recorded according to the PI definitions and final figures produced by the NHS are reviewed for accuracy and queried if they differ from council figures. NHS data had been accurately transferred to council reports.	There is a two month delay between the council's submissions and publication of NHS figures. This is now reported as a snapshot figure in quarterly scorecards as the total number of acute delayed discharges (rather than weekly average).
Customer service statistics:	Email statistics are collated through an entirely manual process by customer service representatives, who follow a defined process. Whilst this may mean	A lot of manual work is involved in the production of email statistics in

Indicator Reviewed	Summary	Weaknesses / Issues requiring attention
a) emails received and handled b) footfall volume per service area c) telephone calls received and handled	there are some inaccuracies in the data responsiveness is more important than complete accuracy and it seems satisfactory for this purpose. Footfall statistics are collated from the ticketing system and a manual count from the visitor log. These had been transferred correctly to the weekly reports. Telephone statistics are collected from data automatically recorded by the system and extracted by reports run according to the PI definitions. These had been transferred correctly to the weekly reports.	particular and this increases the risks of errors. However, there is no obvious way to improve this and the level of accuracy seems reasonable for how these statistics are used.
Number of households for whom positive action has prevented homelessness.	Data is collected from proformas completed by council staff and many other agencies. Training and guidance is given to all caseworkers and the proforma and database used contain DCLG definitions to ensure the PI is reported accurately. There were some discrepancies between the figure reported and the base data provided to audit. Whilst these were not material to the accuracy of the PI they did indicate that there is a lack of rigour in the collection and checking of the figures for this PI.	The PI has been collected by multiple people from multiple agencies for many years but the guidance has never been issued as a reminder. There is no 2nd officer review of figures before they are submitted to the DCLG and no checks undertaken to ensure that the base data is complete.
Timeliness of social care packages	This indicator measures the time from social care assessment to the final day of a social care package and, as such, is not directly useful for determining the timeliness of social care packages. Testing found some small discrepancies, which were likely to have been the result of manual input and have gone undetected because there is no second officer review of the figures. However, this indicator will no longer be collected following the introduction of SALT (short and long term care data collection) in 2014-15 so no recommendation will be made.	Indicator does not appear to be very useful and will no longer be collected in future. No recommendations made regarding this indicator but findings did emphasise value of minimising manual input and ensuring second officer review of figures to prevent and detect any errors.
Number & % of vacant city centre shops	Data is collected from the SX3 revenues system and the accuracy of this PI relies on business rates information. An annual review is done to ensure records are accurate and updates are also made after the annual billing. Testing indicated that the raw data is reliable (with some small differences resulting from updates to records) and the PI has been accurately produced.	There is no documented definition of this indicator or procedure notes on producing the information. The officer who produces this is very experienced and knowledgeable but there would be a risk that a different officer might

Indicator Reviewed	Summary	Weaknesses / Issues requiring attention
		produce the figures differently.
Waste landfilled and recycled (as % of total waste) a) household waste reused, recycled and composted b) municipal waste landfilled.	Data is collected on the ISYS database at waste and recycling sites. It is subject to daily validation checks and discrepancies or inconsistencies are queried. Figures for the indicators are extracted from the database on a weekly basis and entered into a cumulative performance indicator record. Testing found that the indicator had been reported accurately.	One officer is responsible for checking base data, running the weekly reports and producing the final PI figures. There is no review of this process or of the figures by a second officer. The process is heavily reliant on the knowledge and experience of this officer and there are no procedure notes for the checking of data and production of the indicator.
First time entrants to the Youth Justice System (YJS) (per 100,000 population).	Data is collected on the YOTs (Youth Offending Team) database. Recording of data is governed by case recording procedures but checks are also undertaken when the indicator figures are produced. There was a discrepancy between the figures reported for 2013-14 and those re-run for audit. This was a result of errors in case recording being identified and corrected subsequent to the production of the annual figures. Restrictions on data sharing meant figures could not be traced back to source data. However, testing that was undertaken confirmed the PI definitions had been applied correctly. Discrepancies were found between nationally reported Youth Justice Board (YJB) figures and figures collected by the council but explanations for this were considered satisfactory.	There are no procedure notes for producing the indicator and it is heavily reliant of the knowledge and experience of the officer producing them. Regular proactive review of the data gathering process may identify issues before figures are produced.
Number of new affordable homes delivered in York.	Data is collected from planning approvals for homes with an affordable housing requirement. These are recorded on a database and, periodically, the developers are asked to complete a return confirming the number of affordable homes built. Returns are also received from the Homes and Communities Agency on an annual basis and these are added to the council figures. Testing found that affordable homes figures were accurately reported for 2013-14 and the processes in place to identify completions and check this	There are no procedure notes for producing the indicator, it contains numerous manual elements and is heavily reliant of the knowledge and experience of the officer producing them.

Indicator Reviewed	Summary	Weaknesses / Issues requiring attention
	information are reliable.	
Housebuilding: a) net additional homes b) % new homes built on previously developed land	Data is collected from planning permissions for new dwellings. Site visits and correspondence with planning applicants is used to determine when homes have been completed. A definition document details how the indicator should be calculated. Testing did not identify any errors and the officer keeps manual proformas recording the progress for each site and updates these on a completions spreadsheet. For all new homes / sites sampled evidence was available that the home had been completed.	There are many manual elements to collating and producing this indicator that increases the risk of error.
Highways: a) % of roadways that are in poor condition b) % of pathways that are in poor condition	Data for this indicator is collected from the annual condition survey. Ranking of condition relies on the skills, knowledge and experience of the surveyor. Data is transferred electronically to a database and reports are run from this to calculate the indicator. Testing found that all sections of roadways and footways had been rated and the indicator had been produced accurately from this base data.	There are no procedure notes for producing the indicator and reliance is placed on the knowledge and experience of the officer producing the figures.

Audit Opinions and Priorities for Actions

Audit Opinions	
Audit work is based on sampling transactions to test the operation of systems. It cannot guarantee the elimination of fraud or error. Our opinion is based on the risks we identify at the time of the audit. Our overall audit opinion is based on 5 grades of opinion, as set out below.	
Opinion	**Assessment of internal control**
High Assurance	Overall, very good management of risk. An effective control environment appears to be in operation.
Substantial Assurance	Overall, good management of risk with few weaknesses identified. An effective control environment is in operation but there is scope for further improvement in the areas identified.
Reasonable Assurance	Overall, satisfactory management of risk with a number of weaknesses identified. An acceptable control environment is in operation but there are a number of improvements that could be made.
Limited Assurance	Overall, poor management of risk with significant control weaknesses in key areas and major improvements required before an effective control environment will be in operation.
No Assurance	Overall, there is a fundamental failure in control and risks are not being effectively managed. A number of key areas require substantial improvement to protect the system from error and abuse.

Priorities for Actions	
Priority 1	A fundamental system weakness, which presents unacceptable risk to the system objectives and requires urgent attention by management.
Priority 2	A significant system weakness, whose impact or frequency presents risks to the system objectives, which needs to be addressed by management.
Priority 3	The system objectives are not exposed to significant risk, but the issue merits attention by management.

Where information resulting from audit work is made public or is provided to a third party by the client or by Veritau then this must be done on the understanding that any third party will rely on the information at its own risk. Veritau will not owe a duty of care or assume any responsibility towards anyone other than the client in relation to the information supplied. Equally, no third party may assert any rights or bring any claims against Veritau in connection with the information. Where information is provided to a named third party, the third party will keep the information confidential.