



Regulator of
Social Housing

Statistical Data Return (SDR) data quality report 2017-2018

Responsible statistician: Amanda Hall

Version 1.0

October 2018



Introduction

1. In 2015 [the UK Statistics Authority published a regulatory standard for the quality assurance of administrative data](#).¹ The standard was developed in response to concerns about the quality of administrative data and in recognition of the increasing role that such data is playing in the production of official statistics.
2. The Statistical Data Return (SDR) is an administrative data set collected by the Regulator of Social Housing (RSH) and when using this data the RSH follow the standards set in the regulatory standard.
3. This report describes the quality assurance arrangements for SDR data collected by the RSH. It accompanies the SDR statistical release, providing more detail on the regulatory and operational context for data collection and the safeguards that aim to minimise risks to data quality. It is intended to provide added context for users of the data contained in the SDR statistical release and data files.

Regulatory context

4. The regulatory framework for social housing in England provides both the basis for collecting SDR data and the framework which ultimately underpins data quality.
5. The RSH collects SDR data in order to deliver its operational approach to regulating the economic standards set out in the [regulatory framework for social housing in England from the 1st April 2015](#).² The regulatory framework consists of three elements:
 - Regulatory requirements – the requirements with which private registered providers of social housing (PRPs) need to comply
 - Codes of practice – to assist registered providers in understanding how compliance can be achieved
 - Regulatory guidance – further explanatory information on the regulatory requirements, including how the regulator will carry out its role of regulating the requirements.
6. Regulatory requirements include three economic standards for PRPs – Government and Financial Viability Standard, Value for Money Standard, and the Rent Standard. These are the focus of the regulator's operational approach. Regulatory requirements also include those which relate to data and information submission ([information submission requirements](#)).³ As part of the regulatory framework, PRPs are required to submit SDR data by 31 May each year, along with other data returns and regulatory documents at various points during the year.
7. The regulatory framework places the onus for data quality on PRPs and their boards. The [Governance & Financial Viability Standard](#)⁴ sets out the specific

¹ https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/12/images-settingthestandar_tcm97-44370.pdf

² <https://www.gov.uk/government/collections/regulatory-framework-requirements>

³ <https://www.gov.uk/government/publications/information-required-from-registered-providers>

⁴ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/419368/Governance_and_Financial_Viability_Standard_2015.pdf

expectation that 'providers shall communicate with the regulator in an accurate and timely manner'. This expectation is amplified in guidance on the regulatory approach, [Regulating the Standards](#)⁵, which states that the regulator will regard 'the submission of late and incomplete or inaccurate regulatory data may be indicative of a weak control environment' and 'failure to provide accurate and timely data may be reflected in [the regulators'] judgement of a provider's compliance with the regulatory standards'.

8. The RSH acts to address issues of data quality through its stated [regulatory approach](#)⁶. For all providers that own 1,000 or more social housing units, the RSH publishes regulatory judgements of their compliance with the Governance & Financial Viability Standard which includes data quality. Further, the RSH has statutory intervention and enforcement powers ranging from collection of information to removal of PRP officers as set out in [guidance on the regulator's approach to intervention, enforcement and use of powers](#)⁷. The RSH has a duty to exercise its functions in a proportionate way.

Operational context

9. The SDR collects data from all registered PRPs. The first annual collection occurred in 2012 and it has been collected annually from PRPs since. The SDR collects data on stock size, types, location and rents at 31 March each year, and data on sales and acquisitions made between the 12 months up to the 31 March each year. All PRPs are expected to complete the SDR return, with the size of return completed based on the size of PRP stock ownership (those with fewer than 1,000 units of social stock complete a reduced form).
10. The SDR is collected via a web-based data collection system called [NROSH+](#)⁸. The NROSH+ system is hosted and maintained by the RSH (through its digital partner Homes England). It is operated by the RSH's Data Team who control the requirements for data input processes, storage, verification, sign off and extraction of submitted data⁹. Data is either imported or entered into the NROSH+ system directly by PRPs.
11. The RSH sees effective communication with PRPs as an important tool in promoting accurate and timely SDR data submissions. Prior to the start of each financial year, the RSH writes to all PRP Chief Executives (email and hard copy) to highlight regulatory requirements on data submissions including deadlines, the launch of updated data templates and guidance, and expectations on data quality and timeliness in regulatory standards. Regular emails and website communications are used to highlight refreshed SDR data templates and guidance published on the NROSH+ website and to remind PRPs of forthcoming deadlines. In the run-up to the SDR submission deadline of 31 May, the RSH's Data Team proactively contacts providers that have not started their SDR submission according to the NROSH+ system, providing additional support and guidance in completing returns.

⁵

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/698332/Regulating_the_Standards_April_2018.pdf (p19)

⁶ <https://www.gov.uk/government/publications/social-housing-regulation-regulating-the-standards>

⁷ <https://www.gov.uk/government/publications/guidance-on-the-regulators-approach-to-intervention-enforcement-and-use-of-powers>

⁸ <https://nroshplus.regulatorofsocialhousing.org.uk/>

⁹ During 2015/16 the NROSH+ system infrastructure and operation was migrated from an external contractor to HCA-IT. Nodisruption to the delivery of the NROSH+ application was experienced and data collection was unaffected.

Quality assurance processes

12. The RSH's Data Team subject submitted SDR data to a series of internal checks to identify potential quality issues before each individual data return is signed off. The final SDR data file that supports the statistical release is only created from individual SDR returns that are checked and signed off. Where outstanding queries, deemed material to the final dataset, cannot be resolved the data is excluded from the final dataset.¹⁰ In 2017/18 no returns made within the SDR collection period were excluded from the final dataset due to unresolved queries¹¹.
13. There are broadly two types of checks on SDR data submitted by PRPs:
 - Automated validations
 - Manual inspection and sense-checking by RSH (Data Team) staff.
14. Automated validations on SDR data are programmed into the NROSH+ system and check the SDR data at the point of submission. Checks include those that:
 - Ensure every data point is in the correct format (text, numeric, decimal, email, non-blank, etc.).
 - Confirm whether data is consistent, logically possible and within expected limits. For example, automated validations ensure that numbers of units owned are consistent across different parts of the SDR and that local authority areas are reported consistently.
15. Automated validations are either 'hard' or 'soft'.
 - When a 'hard' validation is triggered, the PRP cannot submit their SDR data without the issue being addressed. For example, a hard validation will be triggered when a local authority is present in one question and not in another comparable question.
 - When a 'soft' validation is triggered a warning is given to the PRP to check their information. For example, a soft validation may fire when a value appears to be outside of a normally expected range. PRPs can submit with soft validation errors, but are expected to submit a supporting document to explain the error. Where documents are not submitted or the explanation is insufficient the RSH (Data Team) will investigate further and may contact the provider to discuss the issue.
16. Following submission and automated checks, the RSH (Data Team) runs a systematic programme of manual inspections and sense checking on all submitted data before it is signed off within NROSH+. Manual checks are also completed after SDR data is received. The level of manual checking performed on each return is based on the complexity of the provider and the number of LAs the provider operates in, however all returns are checked for basic consistency and likely errors including where proportions of stock recorded as particular excepted categories is outside of the expected normal

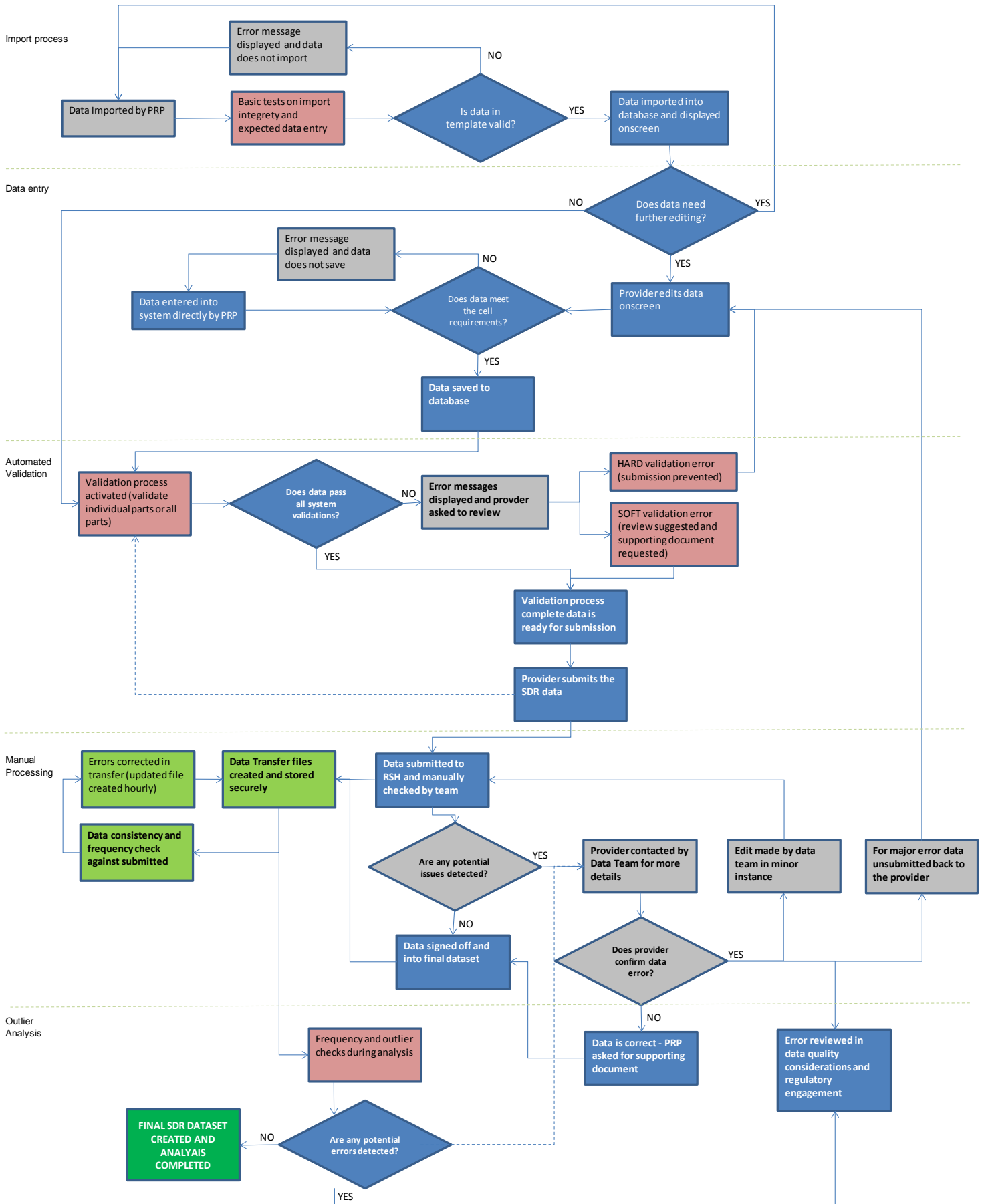
¹⁰ In 2016/17 and 2015/16 there were no returns excluded from the final dataset, but in 2014/15 a total of six returns (all from providers with fewer than 1,000 units) were excluded. Examples of errors in the 2014/15 which led to the return being excluded include: a return where the stock census section was left blank (PRP had stock in 2013/14); the duplication of the number of units entered in multiple sections of the stock census (particularly new entries in non-social housing section); and issues with identification of Affordable Rent stock acquisition route.

¹¹ One return, from a provider with fewer than 1,000 units, was submitted after 31 July 2018 was excluded from the final results as their submission fell after analysis had begun.

range. Areas new to the SDR collection have been further scrutinised to ensure consistency in data where possible.

17. From 2018 the introduction of the stock balance sheet has meant that Large providers are responsible for ensuring acquisitions, losses and movements between categories of stock reported in their SDR account for the change in headline stock counts. This balance sheet has reduced the discrepancies in stock between years, although the RSH considers the information in the balance sheet for each provider and queries any data which appears anomalous (i.e. significant stock movements between categories).
18. Where the RSH's Data Team identifies a potential anomaly with the submitted SDR data, a query is raised with the submitting PRP. The sign off of SDR returns for all providers with 1,000 or more units is dependent on the resolution of all queries. Once a final data set is created no further amendments to the underlying data returns are possible. In 2017/18 all queries were resolved with large PRPs.
19. Almost all the SDR data submitted by PRPs is published at a disaggregated level as part of the statistical release. Releasing data into the public domain serves as an additional route through which erroneous SDR data may be identified by the PRP or third parties. Under the revisions policy described in the statistical release document, where PRPs report errors on submitted data these will be recorded and used to correct data in the subsequent year's statistical release, or through a supplementary release during the year if the level of error is deemed material to the use of the data. The level of revision due to identified errors is documented within the following year's statistical release.
20. If it becomes apparent, either through manual checks or subsequent information, that SDR data has been submitted with significant material errors that may reasonably have been found in a PRP's quality control process, the regulator will consider the extent to which this offers evidence of failure to meet requirements for data quality and timeliness under the Governance & Financial Viability Standard. Within the regulatory framework set out above, the regulator will consider the most appropriate proportionate response to material SDR errors, taking into consideration data quality and timeliness issues across other regulatory data returns.
21. The overall process employed by the RSH Data Team to reviewing the incoming data is outlined in the diagram below.

Figure 1 – Data quality assurance processes.



Coverage error / response rates

22. All PRPs are required to complete the SDR. However, either due to non-submission or exclusion due to unresolved errors there is a small level of known SDR non-response (in 2018 the overall non-response rate was 5%)¹². Following consideration of alternative methods and discussions with the National Statistician's Office and the (then) Department for Communities and Local Government (DCLG) in 2014, weighting has been selected as the preferable solution to account for the small proportion of census respondents who did not submit any data given the incomplete prior year data for some of the non-submitting PRPs and the relative simplicity of the dataset.
23. Using other administrative data held by the RSH it has been identified that to date all of the non-respondents of the SDR data set are small providers, i.e. those which own fewer than 1,000 social housing units/bedspaces. As it is possible to segregate the information supplied by large and small providers in the dataset, weighting has only been applied to the data submitted by small providers (with the exception of the optional SDR questions covering mutual exchanges and evictions). For more information please refer to the weighting section of the statistical release.

Data transfer and data protection

24. Individual SDR returns are stored securely with the NROSH+ infrastructure, accessible to the submitting PRP and RSH staff. The individual returns are also collated into a single data transfer file and are held within a restricted area on the RSH internal server.
25. The data transfer file is subject to a number of checks to ensure consistency with the underlying data. These include checks on PRPs included in each table and spot checks to ensure individual PRP returns are captured correctly.
26. The analysis of the SDR data is undertaken following the UK Statistical Authority Code of Practice for Official Statistics guidelines, ensuring data is held securely and analysis remains accessible only to members of the analysis and production team until the statistics release date. However, the data is utilised for additional specific, internal analysis and administrative purposes as part of our regulatory processes. These analyses are not published or shared beyond key teams within the RSH. Please see pre-release notification list available on the [RSH pages](#).¹³
27. Data submitted by PRPs is redacted within the public release to remove all contact information submitted within the Entity Level Information (ELI) section. This contact information is not publically available.

On-going assessment of data quality

28. The RSH keeps data quality of the SDR under constant review, with reference to the UK Statistical Authority Code of Practice for Official Statistics and supporting documentation. The SDR statistical release provides a commentary of the strengths and weaknesses of statistics presented.

¹² In 2016/17 and 2015/16 the non-response rate was 5% and in 2014/15 it was 4%.

¹³ <https://www.gov.uk/government/collections/statistical-data-return-statistical-releases>

29. The UK Statistical Authority Administrative Data Quality Assurance Toolkit states that the assessment of the quality assurance level should be pragmatic and proportionate. There are four levels of assurance that can be assigned to an administrative data source determined through the use of risk/profile matrix.
30. With respect to the UK Statistical Authority Administrative Data Quality Assurance Toolkit and in the context of other statistics published by UK Government and its agencies, the RSH has determined that the SDR is likely to be categorised as 'low quality risk – medium public interest' (with a requirement for basic/enhanced assurance). Notwithstanding this, the RSH aims for the highest standards of data quality possible within the constraints of available resources and the existing regulatory and operational context. Through on-going internal analysis, the RSH seeks to understand SDR data quality and identify potential means by which it may be improved.
31. The designated responsible statistician for the SDR statistical release was Amanda Hall.
32. For further information please contact the Referrals and Regulatory Enquiries Team on 0300 1245 225 or email at enquiries@rsh.gov.uk



© RSH copyright 2018

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated.

To view this licence, visit: nationalarchives.gov.uk/doc/open-government-licence/version/3

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at: <https://www.gov.uk/rsh>.

Any enquiries regarding this publication should be sent to us via

Email: enquiries@rsh.gov.uk

Telephone: 0300 124 5225.

Or write to:

Regulator of Social Housing

1st Floor - Lateral

8 City Walk

Leeds

LS11 9AT

The Regulator of Social Housing regulates registered providers of social housing to promote a viable, efficient and well-governed social housing sector able to deliver homes that meet a range of needs.