



Regulator of
Social Housing



Data quality and methodology

Private registered providers
social housing stock and rents
in England statistics

2020-2021

Contact information

Responsible statistician: Amanda Hall

Public enquiries: enquiries@rsh.gov.uk or 0300 124 5225

OFFICIAL

Contents

Contents.....	1
Introduction	2
Regulatory context	5
SDR collection.....	7
Statistical release methodology.....	10
Why not have your say on our statistics in 2021/22?	15
Annex A: Quality assurance processes - process map	16

Introduction

This report describes the quality assurance arrangements for the private registered provider (PRP) social housing stock and rents in England statistics, providing more detail on the regulatory and operational context for data collection and the safeguards that aim to maximise data quality.

Background

The statistics we publish are based on data collected directly from PRPs through the Statistical Data Return (SDR) survey. We use the SDR data extensively as a source of administrative data within the delivery of our operational approach to regulating the economic standards (see regulatory context on page 5). The United Kingdom Statistics Authority (UKSA) encourages public bodies to use administrative data for statistical purposes, as such, we publish these data.

National Statistics status

The statistics derived from the SDR data and published as private registered provider social housing stock and rents in England are considered by the Office for Statistics Regulation (OSR), the regulatory arm of the UKSA, to have met the highest standards of trustworthiness, quality and public value.

Designation history

These statistics were designated as a National Statistic by the OSR in 2014 following an assessment against the [Code of Practice for Statistics](#). Following the designation as a National Statistic, producers of the statistics must comply with the statutory requirement to ensure the Code of Practice continues to be observed. We keep the trustworthiness, quality and value of the statistics under constant review.

Publication schedule

These statistics are usually fully published in the Autumn each year, with the data pre-announced in the release calendar.

All data, supplementary tables, data tools and additional information (including a list of individuals (if any) with 24hour pre-release access) are published on our [statistics pages](#).

These statistics are also presented in the Department for Levelling Up, Housing and Communities (DLUHC), formerly the Ministry for Housing, Communities and Local Government (MHCLG) live tables and our registered provider social housing statistics.

Improvements since designation

Since being designated a National Statistic, the following improvements have been made:

Accessibility

Excel based “*look up tools*”, allowing the easy interrogation of data at a PRP and local authority (LA) level are created annually.

Additional supplementary notes, guidance and documents to provide simple access to information to understand the statistics in greater detail (including this data quality and methodology note).

In line with the [Web Content Accessibility Guidelines](#) (WCAG) we have sought to make more of our publication accessible to users by introducing html versions (or html summaries) of documents.

In 2021 we ensured that our additional table and raw data files were published in spreadsheets which employed the principles of accessibility as recommended in the Government Statistical Service guidance.

Timeliness

We keep the publication timescales under close review. In 2021 we pushed back the publication date from the end of September to the end of October. This allowed for the simultaneous publication of PRP and LARP statistics (following the first joint publication in the pandemic delayed 2020 release, which saw the PRP statistics released in February 2021 rather than September 2020). While later than in previous years (2019 and earlier), we believe this October publication date still provides timely access to our data for our key users.

Transparency

The revisions policy and processes have been enhanced to improve transparency in data changes.

Relevance to users

A briefing note style was adopted following user feedback on use and value.

A ‘quick feedback’ link was introduced across all published documents to allow users an easy facility to provide feedback.

User feedback

We are always keen to increase the understanding of the data, including the accuracy and reliability, and the value to users. Please email feedback, including suggestions for improvements, or queries as to the source data or processing to enquiries@rsh.gov.uk.

Quality assurance of administrative data

The data used in the production of these statistics are classed as administrative data. As part of our compliance to the Code of Practice, and in the context of other statistics published by the UK Government and its agencies, we have determined that the statistics drawn from the SDR are likely to be categorised as low quality risk – medium public interest (with a requirement for basic/ enhanced assurance).

The publication of these statistics can be considered as medium public interest, as there has been mainstream media interest, but they have only moderate economic and/or political sensitivity. Concerns over data quality are considered low given the data checks by providers and our data quality checks conducted on the submitted data and analytical processes.

Notwithstanding this, we aim 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, we seek to understand the strengths and limitations of the data, the overall quality of the data and to identify potential means by which it may be improved.

Regulatory context

The regulatory framework for social housing in England provides both the basis for collecting SDR data and the framework which ultimately underpins data quality. For more information about the Regulator of Social Housing and the regulatory framework please see [our website](#).

Regulatory framework

We collect SDR data to facilitate our operational approach to regulating the economic standards set out in the [regulatory framework for social housing in England](#). The regulatory framework consists of three elements:

- **Regulatory requirements** – the requirements with which PRPs need to comply (including our three economic standards (Governance and Financial Viability, Value for Money and Rent Standard) and requirements on data and information submission).
- **Codes of practice** – to assist PRPs in understanding how compliance can be achieved.
- **Regulatory guidance** – further explanatory information on the regulatory requirements, including how we will carry out our role of regulating the requirements.

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.

The regulatory framework and data quality

The regulatory framework places the onus for data quality on PRPs and their boards. The [Governance and Financial Viability Standard](#) sets out the specific 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](#). This states that we will consider that “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 regulator’s] judgement of a provider’s compliance with the regulatory standards”.

Addressing issues with data quality

We have a range of statutory and enforcement powers and act proportionately to address issues of data quality through our [regulatory approach](#). We publish regulatory judgements relating to compliance with the Governance and Financial Viability Standard (which includes data quality) for PRPs that own 1,000 or more social housing units.

Governance of data and statistics at RSH

The statistician responsible for the publication of these statistics is also responsible for the SDR data collection and the cleansing of incoming SDR data; working with PRPs to directly address anomalies within the data submissions and producing the final data set on which the statistics are produced.

All SDR data are stored and analysed within password protected government secure networks and access to the sector level analysis work undertaken on the data is restricted until after publication (PRP level data is accessed by our staff as part of operational work). Further information on the data quality assurance processes we employ is provided on page 7.

Data submitted by PRPs is redacted within the public release to remove all contact information submitted within the Entity Level Information section. This contact information is not publicly available. There is no other administrative data held by us which can be made available for use in statistics. However, we publish a range of summary data from other information collected which are available on our website's [analysis and statistical reports](#) page.

SDR collection

All PRPs are expected to complete the SDR. The size of return completed is based on the size of the PRP, with those owning 1,000 or more units of social stock providing more information.

History

The first collection occurred in 2012 and it has been collected annually from PRPs since. The SDR collects data on stock size, types, location and rents as at 31 March each year, and data on sales and acquisitions made during the 12 months up to the 31 March each year.

Systems

The SDR is collected via a web-based system called [NROSH+](#). The NROSH+ system is hosted and maintained by us (through our digital partner Homes England). We control the requirements for data input processes, storage, verification, sign off and extraction of submitted data and produce the statistical releases. Data is either imported or entered directly into the NROSH+ system directly by PRPs.

We are currently considering our data collection needs and are looking to redevelop our data collection system.

Communication with data suppliers

We work closely with PRPs, through email messages and phone discussions, to ensure there is a common understanding of the data collection requirements throughout the data collection process. Guidance materials are also promoted to users and published on NROSH+.

Quality assurance processes

We do not have oversight of the systems and data quality assurance processes employed by the PRPs before submitting data to SDR. However, we do provide clear guidance and documentation on the NROSH+ system and subject SDR submissions to a series of checks to identify potential quality issues before each data return is signed off.

The final SDR data file that supports the statistical release is only created once all outstanding queries are resolved. Any returns not meeting our quality standards are excluded from the final dataset.

Submission checks

SDR data submitted to us is subject to both automated validation checks and manual inspection.

Automated validations are programmed into NROSH+ and check the SDR data at the point of submission for correct formatting, consistency and logical possibility (within expected limits). For example, ensuring numbers of units are consistent across different parts of the SDR and that LA areas are reported consistently. Automated validations are either 'hard' or 'soft'.

Hard validations – PRPs cannot submit without the issue being resolved (e.g. when an LA area is present in one question and not in another comparable question).

Soft validations – PRPs can submit but are required to check their information, and if correct to submit a supporting document (e.g. when a value appears to be outside of a normally expected range, such as higher than expected rents).

Manual inspections are systematically undertaken on all data submitted. All returns are checked for basic consistency and likely errors, for example where proportions of stock recorded as particular excepted categories are outside the expected range or where rents are particularly high or low. This includes comparison to previously submitted data with unexpected movements in stock levels or changes outside those anticipated in reported rent values being queried with PRPs. Where we identify a potential anomaly with the submitted SDR data, a query is raised with the submitting PRP. The sign-off of SDR returns is dependent on the resolution of all queries which could materially impact the quality of the published data. This overall checking process is outlined in the diagram in Annex A.

Post publication checks

SDR data (excluding contact details or optional pilot year questions), are 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 data may be identified by the PRP or third parties.

Misreporting

There are no numerical measures of misreporting of SDR data by PRPs. However, one source of possible quality weakness is inconsistent interpretation of guidance with providers not applying this consistently across the sector. This issue is most likely to arise where there are technical or legal definitions that are complex or, to some degree, ambiguous. It is more likely to arise among PRPs owning fewer than 1,000 units due to the reduced level of contact with regulation staff and their specialisms in certain types of activities (e.g. supported housing/ leasehold). Please see **technical notes and definitions** for more information on other factors which impact on the data collected.

Corrections

Where errors in the SDR data are discovered within a survey year, either through regulatory activity or through provider contact, we allow providers to resubmit SDR data through the NROSH+ portal. Returns can be amended until mid-March the year following their launch. However, it should be noted that due to the collection of the 2020 SDR being delayed as part of our coronavirus response, the opportunity for 2020 SDR revisions was more limited and no 2020 SDR resubmissions were made.

Under the revisions policy (see **technical notes and definitions**), errors identified will be investigated and revised data gathered. Some corrections may be only minor changes to the data, with little or no impact on the published statistics. These changes will be published at the next scheduled release with no specific announcement.

However, if we become aware of substantial errors in the submitted data, statistical process or other methodology and where a major revision to the published data is required, a non-scheduled revision of the statistical release will be published. This will include full details of the revisions, clearly marked data amendments and summary tables showing the overall impact of the changes.

We seek to ensure transparency in processes to maximise user confidence in the quality of our statistical releases.

Data quality regulation

If, through either manual checks or subsequent information, SDR data has been submitted with significant material errors, that may reasonably have been found by a PRP during their internal quality control process, we will consider the extent to which this offers evidence of failure to meet requirements for data quality and timeliness under the Governance and Financial Viability Standard. Within the regulatory framework set out on page 5, we will consider the most appropriate response.

2021 responses

All PRPs are required to complete the SDR. However, due to non-submission or exclusion due to unresolved errors there is a small level of known SDR non-response. In 2021 the overall non-response rate was 4.7%. This is similar to that seen in previous years. Using other administrative data held by us, we have identified that all the PRPs excluded from the SDR data are those which own fewer than 1,000 social housing units/ bedspaces. The response rate for these 'small' providers in 2021 was 94.2%. Data is weighted to account for this small proportion of census respondents for which data is not available (see accounting for missing data section below).

Statistical release methodology

Accounting for missing data

A list of late/ missing returns has been published alongside the release since 2018.

In 2014, following consideration of alternative methods and discussions with the National Statistician's Office and (now) DLUHC, weighting was selected to account for the small proportion of missing data. This method was chosen given the incomplete prior year data for some of the non-submitting PRPs and the relative simplicity of the dataset.

Weighting

SDR data (2012 to 2021) has been weighted. Weighted data is highlighted in the release and is marked in relevant tables by the use of an uppercase 'I'. Data are weighted in the following categories:

Categories	Large PRPs	Small PRPs
General needs		✓
Supported housing		✓
Housing for older people		✓
LCHO		✓
Social leasehold		✓
Non-social rented/ leasehold		✓
Affordable Rent (all categories)	not weighted	not weighted
Evictions (all categories)	✓(some ¹)	N/A
Mutual exchanges ²	✓	N/A
No. of Decent Home Standard failures		✓

¹ Excepting 2012 data, where an accurate response rate for this question cannot be determined and from 2020 where this question became mandatory for all large PRPs.

² Excepting 2012 data, where an accurate response rate for this question cannot be determined.

Caution should be used when viewing weighted results for evictions and mutual exchanges. Given the optional nature of the questions (in some or all years), non-responding PRPs may not have the same trends as responding PRPs. It may be that those with high rates of evictions

were more likely to choose not to disclose that information in an optional question or those with a low rate may have felt it unnecessary.

The figures presented for 2020 evictions provided some evidence that our prior weighting was likely to have provided a more accurate total number of evictions. Removing data from the 2020 dataset for those providers who, in 2019, did not respond to the evictions question resulted in a recording of 7,428 evictions. Weighting this figure using the methodology employed in previous years saw a weighted figure of 9,653 evictions which was closer to the actual recorded evictions in 2020 (10,311) than the unweighted figure of 7,428. However, it was not possible to confirm if the increase seen in 2020 was due to the presence of previously missing data or a true rise in the number of evictions.

2021 weighting

The impact of weighting the 2021 data is shown in published supplementary tables. As in previous years the effect of weighting on totals is relatively minor (0.2% to 2.0%) given the minimal missing information.

Basic method

Weighting based on the response rate to the SDR has been applied to the categories shown above. It assumes that the trends in the data submitted by PRPs would also apply to the PRPs who did not submit (e.g. that the proportion of social housing stock owned in each region of England is the same for the small providers that did not respond as it is for the small providers that did).

The general formula used to conduct this weighting was as shown:

$$\text{weighted result} = \text{unweighted result} \times \frac{1}{\text{response rate}}$$

For regional and national totals, weighting was carried out at the LA level and aggregated upwards. Due to the discrete nature of the data (number of units), all data were rounded up (i.e. a weighted result of 10.1 units would be recorded as 11 units as it represented a figure greater than ten units).

Exceptions

The number of Affordable Rent units was not weighted. To own this type of stock, PRPs must be signed up to the Affordable Homes Programme which requires additional reporting and active engagement with Homes England/ Greater London Authority and with us. The likelihood that any Affordable Rent stock has not been captured in the SDR is therefore considered to be very small. Accordingly, these units were removed from the dataset before weighting occurred, then added back in.

Where Affordable Rent data could not be split by stock type; for example, the number of supported housing and housing for older people it was assumed that the distribution of Affordable Rent stock followed the distribution of social rent stock (e.g. supported housing and housing for older people Affordable Rent was assumed to be divided between the two component stock types in the same proportion as the units not designated as Affordable Rent supported housing and housing for older people stock.

Average rent and service charge calculations

General needs, supported housing and housing for older people rental figures in this release are calculated using data supplied by large PRPs only, as small providers are not required to submit detailed rental data. All PRPs with Affordable Rent stock are required to submit rental information for that stock regardless of the total number of units owned.

Calculation of averages

All averages relating to rents in this statistical release are fully weighted by stock owned by PRPs for the appropriate geography and/ or sub-group.

Average service charges and gross rents

The average service charges presented in the rent sections relate only to the stock where there is a 'housing benefit eligible' service charge present. Therefore, zero service charges are excluded from this calculation.

However, gross rents presented in these tables do include stock without a service charge. Because of this, the sum of the average net rent and average service charge will not equal the average gross rent.

Calculation of social rent rates

PRPs are required to follow the [guidance we set out when calculating social rent rates](#). It should be noted that social rent rates are not applicable to homes let under the Affordable Rent programme, classified as temporary social housing or intermediate rent properties.

Unit sizes for which rent data are collected

PRPs owning 1,000 or more units are required to submit LA level breakdowns for rent and service charges for the following unit sizes. PRPs owning fewer than 1,000 units submit rent figures at a PRP level only, this combines all unit sizes and locations. Analysis presented in the statistical release focuses on the detailed rent data submitted by larger PRPs only.

General needs and Affordable Rent general needs	Supported housing/ housing for older people and Affordable Rent supported housing/ housing for older people
Bedspaces/ non-self-contained	Bedspaces/ non-self-contained
Bedsit	Bedsit
1 bedroom	1 bedroom
2 bedroom	2 bedroom
3 bedroom	3 bedroom
4 bedroom	4 or more bedrooms
5 bedroom	
6 or more bedrooms	

For all material in the 2021 release (including briefing notes, supplementary tables and the 2021 data set) visit our [statistics pages](#).

Quality assurance of the published statistics

The data, briefing note, look up tools and tables are quality assured by analysts within our statistics production team. This process ensures the figures are consistent across the release, and match the raw data submitted through SDR. Each check being signed off and recorded by the responsible statistician when it has been completed.

Revisions

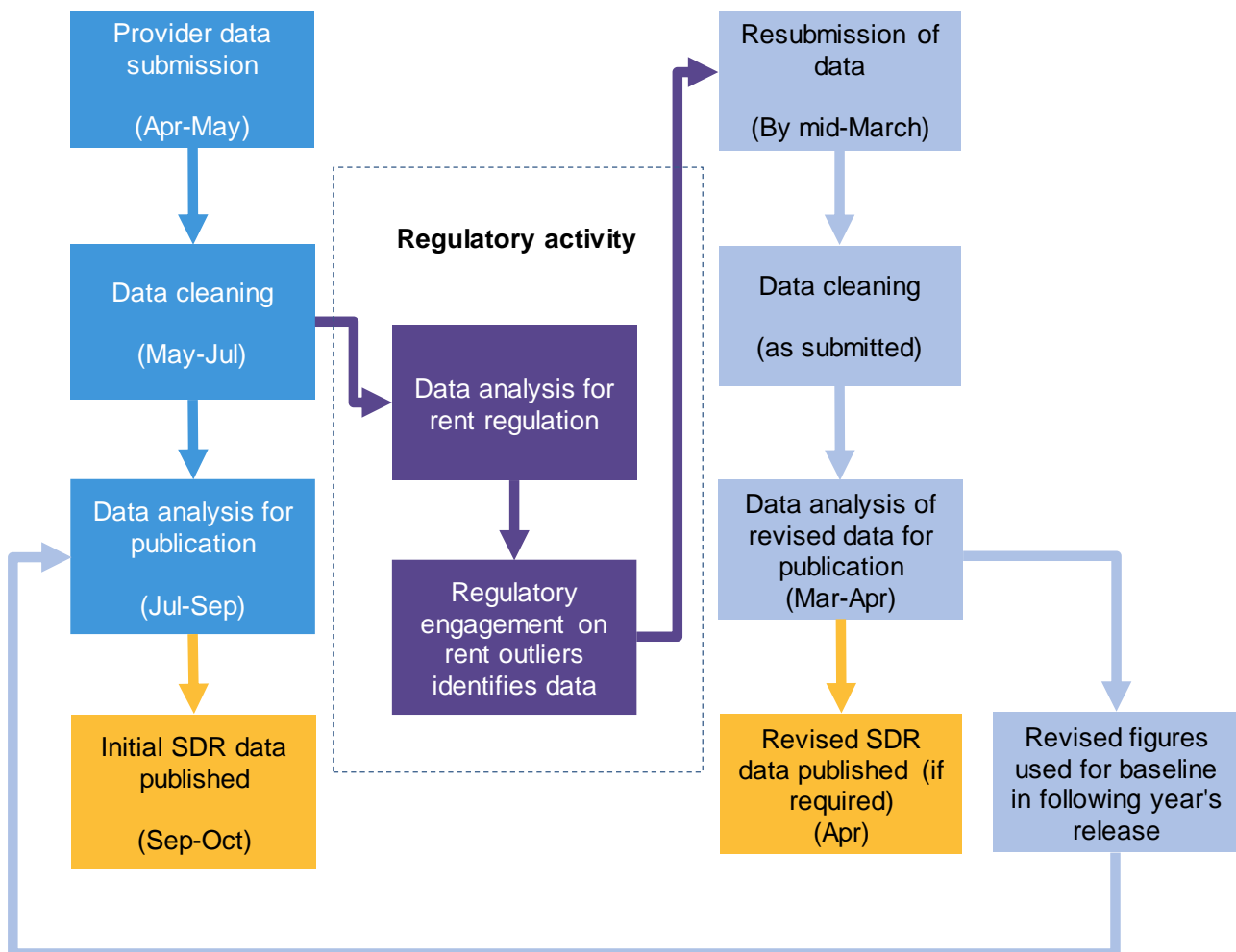
Under the revisions policy (see **technical notes and definitions**), errors identified will be investigated and revised data gathered. Some corrections may be only minor changes to the data, with little or no impact on the published statistics. These changes will be published at the next scheduled release with no specific announcement.

However, if we become aware of substantial errors in the submitted data, statistical process or other methodology and where a major revision to the published data is required a non-scheduled revision of the statistical release will be published. This will include full details of the revisions, clearly marked data amendments and summary tables showing the overall impact of the changes.

Annual cycle of regulatory activity

The SDR data is used to inform our engagement on registered providers' compliance with the Rent Standard. As part of their response to any issues raised, providers subject their data to increased validation and may identify errors in the data submitted. We are committed to ensuring the quality of the SDR data and will gather corrected data from PRPs as part of this work.

We will republish these statistics in the April of the year following the initial publication if the aggregate changes made by providers require a major revision. If a major revision to published data is not required, the changes will be incorporated (and clearly marked) in the published baseline data for the following years' release.



Why not have your say on our statistics in 2021/22?

We want to hear your views on how the format and range of documents in this statistical release meet your needs. Please email feedback, including suggestions for improvements to enquiries@rsh.gov.uk or click below to quickly rate how this document meets your needs.



All needs met



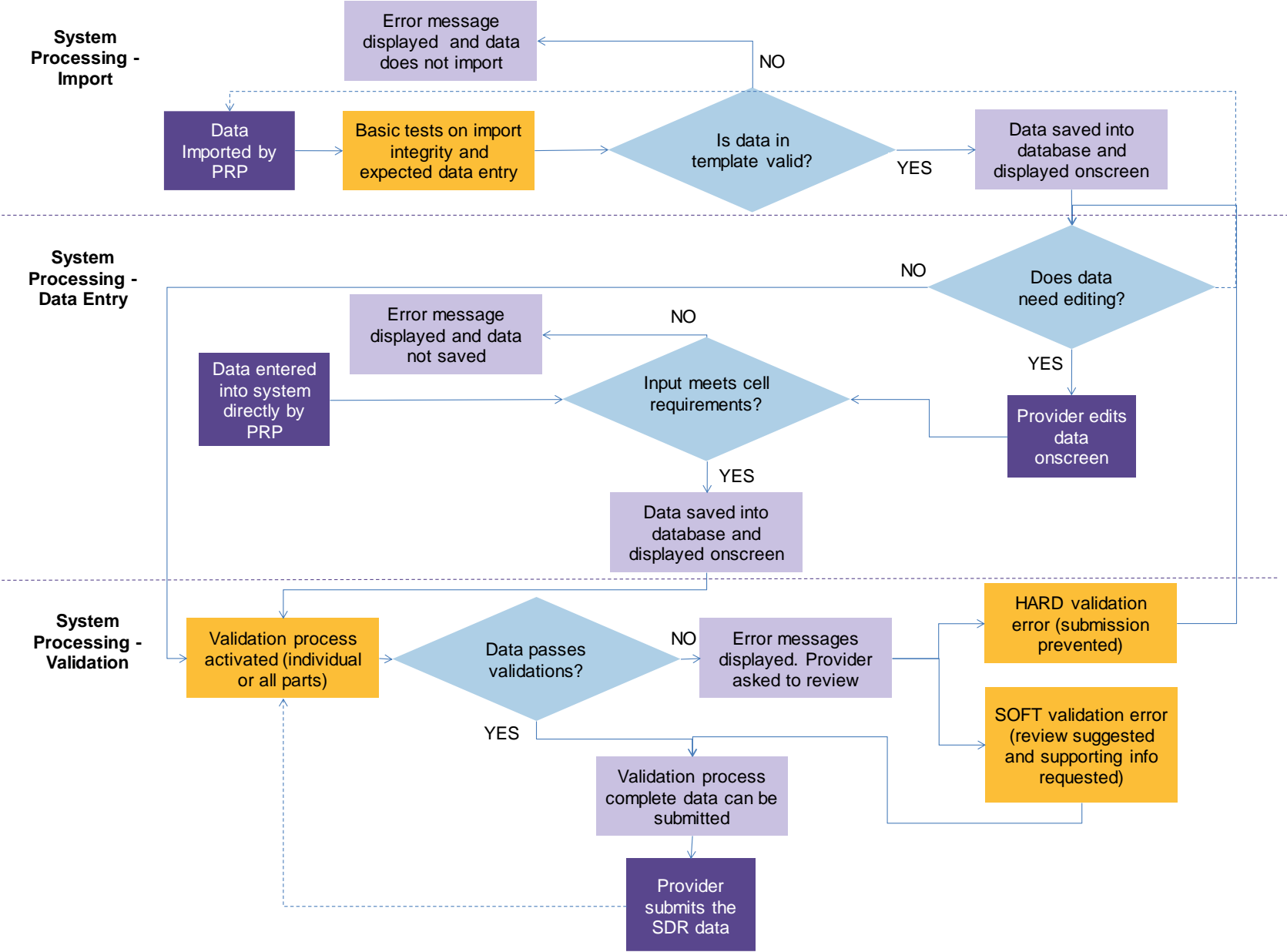
**Some needs
met**



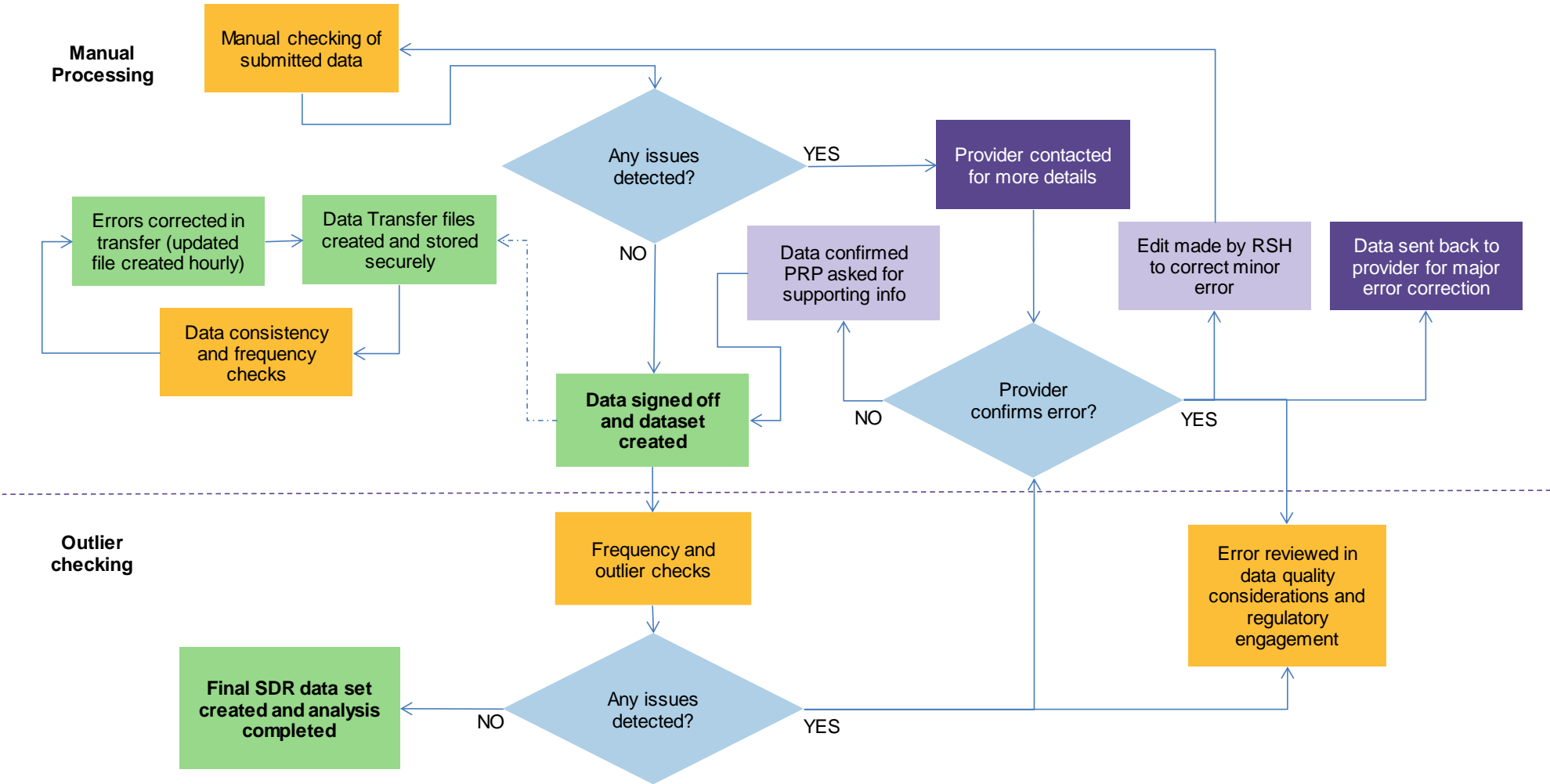
No needs met

Annex A: Quality assurance processes - process map

Data import, entry and system validation (pre-submission to RSH)



Manual data processing and outlier checking (post submission to RSH)





© RSH copyright 2021

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: www.gov.uk/rsh

Any enquiries regarding this publication should be sent to us via enquiries@rsh.gov.uk or call 0300 124 5225.

or write to:

Regulator of Social Housing
Level 2
7-8 Wellington Place
Leeds LS1 4AP

RSH regulates private 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.