# Data quality and methodology

## Local authority registered provider social housing stock and rents in England statistics

### 2022-2023

**Contact information**

Responsible statistician:     Amanda Hall

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

# Contents

# Introduction

This report describes the quality assurance arrangements for the local authority registered provider (LARP) 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 LARPs through the local authority data return (LADR) survey. We use the LADR data extensively as a source of administrative data. We first collected these data in 2020 and use them to deliver our operational approach to regulating rents (see regulatory context on page 6). The United Kingdom Statistics Authority (UKSA) encourages public bodies to use administrative data for statistical purposes and, as such, we publish these data.

The first publication of these data as collected through the LADR was in February 2021 (relating to data as at 31 March 2020). In previous years, data on owned local authority stock and rents were collected through the local authority housing statistics (LAHS) and published by the Department for Levelling Up, Housing and Communities (DLUHC), formerly the Ministry for Housing, Communities and Local Government (MHCLG), in the corresponding statistical release. For more information see Local authority housing data - GOV.UK (www.gov.uk).

## National Statistics status

The statistics derived from the LADR data and published as LARP 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

During 2020/21 we asked the OSR to review our planned publication of these statistics following the transfer of responsibility for the data collection and production from the, then, MHCLG to the Regulator of Social Housing (RSH). The OSR compliance check focused particularly on our work to maintain and enhance the quality of the statistics through this transition and in February 2021 they confirmed that these statistics could continue to be designated as National Statistics.

The OSR reported a range of positive features that demonstrated the trustworthiness, quality, and value of these statistics. These are outlined in detail in the 2020 data quality and methodology note.

**User feedback**

We are 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.

**Publication schedule**

We intend to publish these statistics in 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 24 hour pre-release access) are published on our statistics pages.

These statistics are also presented in DLUHC live tables and our registered provider social housing statistics.

## Quality assurance of administrative data

The data used in the production of these statistics are classed as administrative data. In 2015 the UKSA published a regulatory standard for the quality assurance of 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 LADR 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. This categorisation is the same as for the previous LAHS and current statistical data return (SDR) data based statistical releases, both of which have similar collection and analysis methodologies and coverage.

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 ongoing 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 underpins the collection of LADR data. For more information about the RSH and the regulatory framework please see our website.

## Regulatory standards

First collected in 2020, the LADR data are used to facilitate our operational approach to regulating the Rent Standard.

Paragraph 3.19 of the rent standard requires local authorities registered with us to *"communicate with the Regulator in an accurate and timely manner. This includes all data and information required by the Regulator in respect of compliance with this Standard."* This requirement includes the annual submission of LADR data.

## The regulatory framework and data quality

In Regulating the Standards, we set out the operational approach to regulation and emphasise the importance of providers supplying timely and accurate data. More detail on our approach and expectations regarding data and information submission from LARPs are set out in paras 3.6 to 3.10 of the document.

## Governance of data and statistics at RSH

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

All LADR 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 (LARP level data are accessed by our staff as part of operational work). Further information on the data quality assurance processes we employ is provided on page 8.

Data submitted by LARPs 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 nor local authority specific data collected by us which can be made available for use in statistics.

# LADR collection

All local authorities registered with RSH (LARPs) are expected to complete the LADR. The return provides us with the data required for the risk-based, data-driven approach to regulating local authority compliance with the rent standard.

## History

The LADR was collected for the first time in 2020. This followed a pilot collection in 2019, where we worked with local authorities to encourage engagement and to identify areas where our systems and guidance could be improved. From 1 April 2020 the LADR was included on the single data list which outlines datasets local government must submit to central government.

In 2022 we discontinued the collection of dwelling equivalent data following a process of user feedback gathering and review. Details of this change, including the impact assessment, can be found in our 2021 publication.

### Coverage

LADR does not collect data from all local authorities, only those registered with us as providers of social housing. However, as set out in section 114A of the Housing and Regeneration Act 2008, all local authorities who provide social housing are subject to compulsory registration. Therefore, the data collected in LADR can be considered a complete census of local authority owned[1] social housing. Where the managing agent of local authority stock – the arm's length management organisation (ALMO) – also owns units, the ALMO will be registered as a private registered provider for these units. The ALMO will complete the SDR providing details of the stock they own and record LARP owned stock as units managed on behalf of others.

## Comparison to previous data collection (LAHS)

The LAHS collection, collected by DLUHC, includes a wide range of information related to local authority housing from all lower tier local authorities (including unitary authorities). Data are collected as a snapshot as at 31 March and for the whole of the financial year.

In 2020, following discussions with local authorities, a number of questions relating to social housing stock numbers and rents were transferred from LAHS to LADR in order that we could collect information essential for regulating rents and to ensure no duplication (nor increased

---

[1] Local Authorities are asked to report units they own, and which are not leased to other registered providers. Where units are leased (on a lease of any length) to a private registered provider that body is considered to be the owner. The definition of ownership focuses on the body with the direct landlord relationship with the tenant.

burden) in data supply. The 2020 data quality and methodology note provides more information on the transfer of questions.

The discontinuing of the dwelling equivalent count for non-self-contained units in 2022 will mean that the annual comparisons of non-self-contained units can no longer be made between LAHS and LADR.

For more information on the changes between the two collections and the impact these may have on these statistics please see the **technical notes and definitions.**

## Systems

The LADR is collected via a web-based system called NROSH+. We control the requirements for data input processes, storage, verification, sign-off and extraction of submitted data and produce the statistical releases. Data are either imported or entered directly into the NROSH+ system by LARPs.

In 2023 a revised NROSH+ application was launched. The principles of the site remained the same, but new features such as on-screen live validation updates and better navigation were introduced. We do not believe the new system had any impact on providers and the response rates to the survey have remained similar to previous years.

## Communication with data suppliers

We work closely with local authorities, 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 promoted and published on NROSH+ and we work closely with the central local information partnership – housing (CLIP-H) to develop questions and guidance.

## Quality assurance processes

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

The final LADR data file that supports the statistical release is only created once all outstanding queries which could materially impact the quality of the published data are resolved. Any returns not meeting our quality standards are excluded from the final dataset.

A list of all LARPs who have not submitted data or whose data did not meet our quality standards will be published alongside the data in the years this is required.

## Submission checks

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

**Automated validations** are programmed into NROSH+ and check the LADR 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 LADR. Automated validations are either 'hard' or 'soft'.

> **Hard validations** – LARPs cannot submit without the issue being resolved (for example, when data on stock numbers is missing but a rental value is given).

> **Soft validations** – LARPs can submit but are required to check their information, and if correct to submit a supporting document (for example, 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. Returns are checked for basic consistency and likely errors, for example where proportions of stock are 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 LARPs.

Where we identify a potential anomaly with the submitted LADR data, a query is raised with the submitting LARP. This overall checking process is outlined in the diagram in Annex A.

We also use the provisional LAHS headline stock data each year to assess the quality and coverage of the LADR data. We review individual LA discrepancies against queries raised during cleaning to identify reasons for these. Where we do not have information, we contact LARPs to confirm the data submitted to LADR were accurate. Data from LADR is provided back to DLUHC to assist with further data checks on the LAHS data during the finalisation of the dataset. For more information on the differences between LADR and LAHS and the impact this may have on these statistics please see **technical notes and definitions.**

## Submission checks and sign-off

It should be noted that the process of signing off data is distinct from our ongoing regulatory work. The sign-off of data confirms that we have investigated areas of potential data error, as highlighted in the validation and checking work, and accepted that the provider has submitted data they believe to be accurate. However, during the course of our regulatory activity, the data will be reviewed alongside other evidence, and we may subsequently challenge classification of stock, rent setting or any other aspect covered by the data with providers. As such, the sign-off and publishing of these data does not constitute our agreement that the provider has appropriately classified units or complied with our standards.

## Post publication checks

LADR data (excluding contact details) 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 LARP or third parties.

## Misreporting

There are no numerical measures of misreporting of LADR data by LARPs. 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. Please see **technical notes and definitions** for more information on other factors which impact on the data collected.

## Corrections

Where errors in the LADR data are discovered within a survey year, either through regulatory activity or through provider contact, we allow providers to resubmit LADR data through the NROSH+ portal. Returns can be amended until mid-March the year following their launch.

# Statistical release methodology

The data presented in the LARP social housing stock and rents statistics are drawn from the LADR data.

## Accounting for missing data

All questions in the LADR are mandatory and all LARPs are expected to submit the return.

**2023 responses**

As in 2021 and 2022, the 2023 response rate was 100%.

Due to the 100% response rate and all data queries being resolved in 2023 we have not needed to impute any figures to account for missing or suspected inaccurate data.

## 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 LADR. Each check is 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.

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

### Annual cycle of regulatory activity

The LADR data are collected to inform engagement on registered providers' compliance with the rent standard. As part of their response to any issues raised by us, providers subject their

data to increased validation and may identify errors in the data submitted. We are committed to ensuring the quality of the LADR data and will gather corrected data from LARPs 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 year's release.
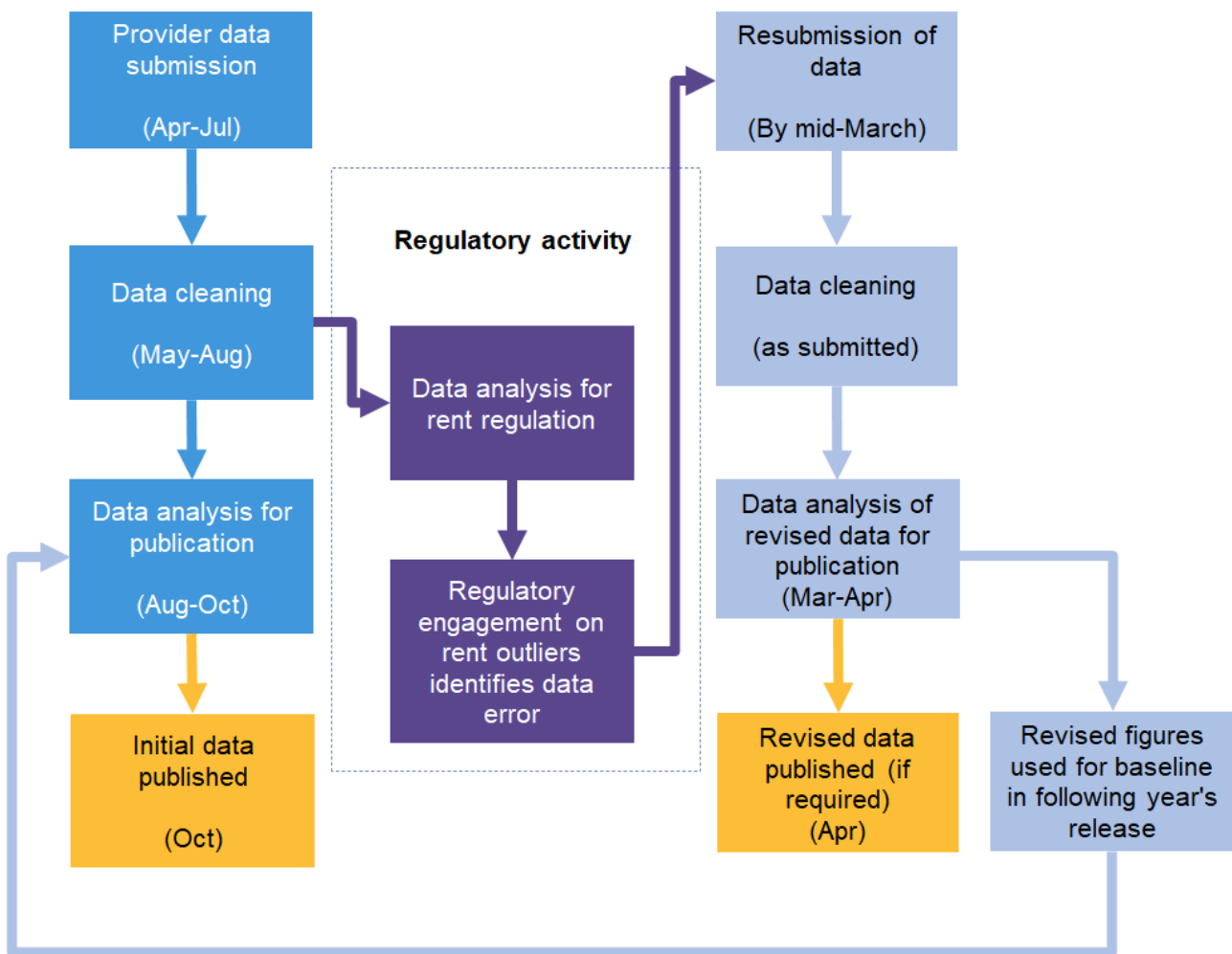
# Why not have your say on our statistics in 2023/24?

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.



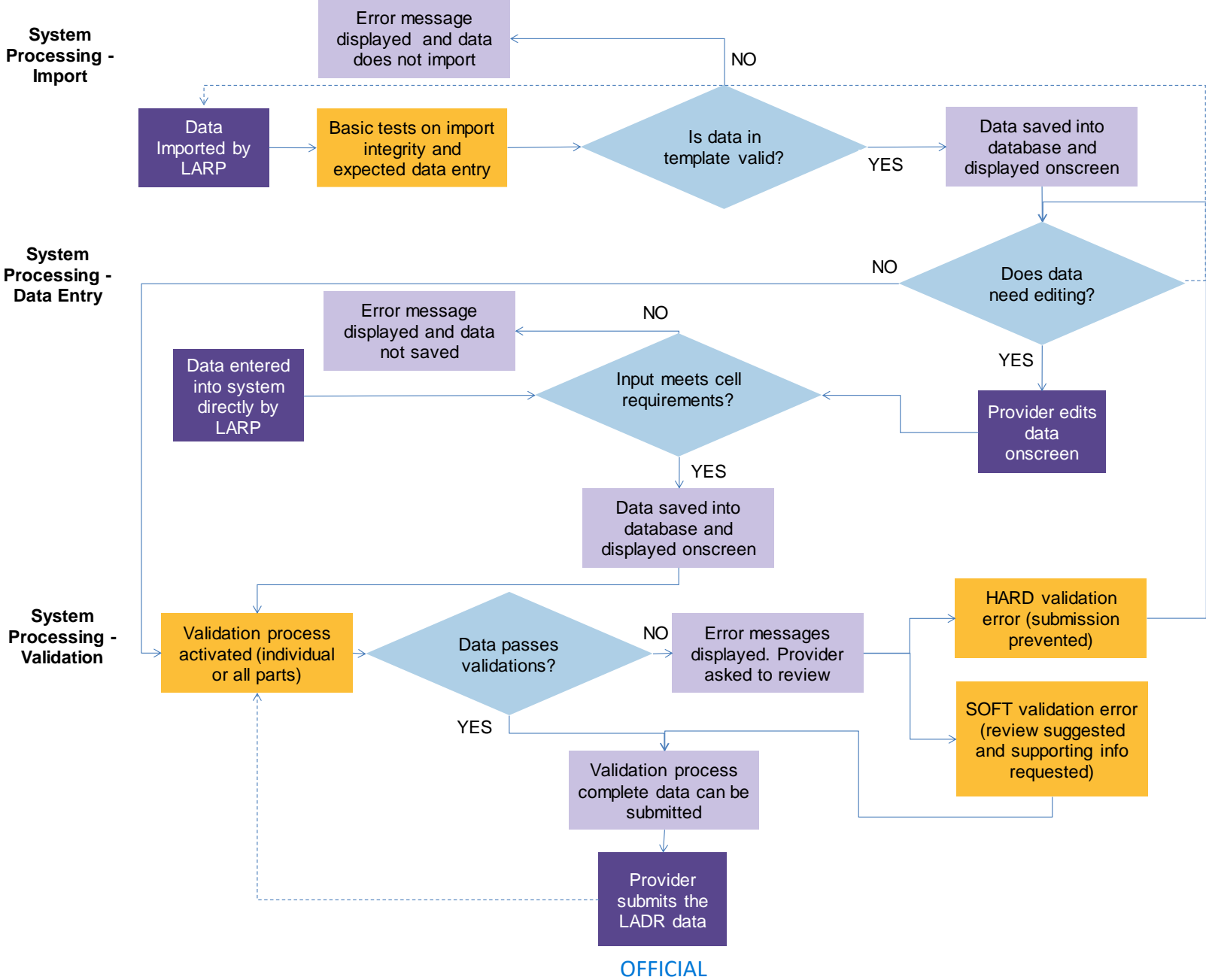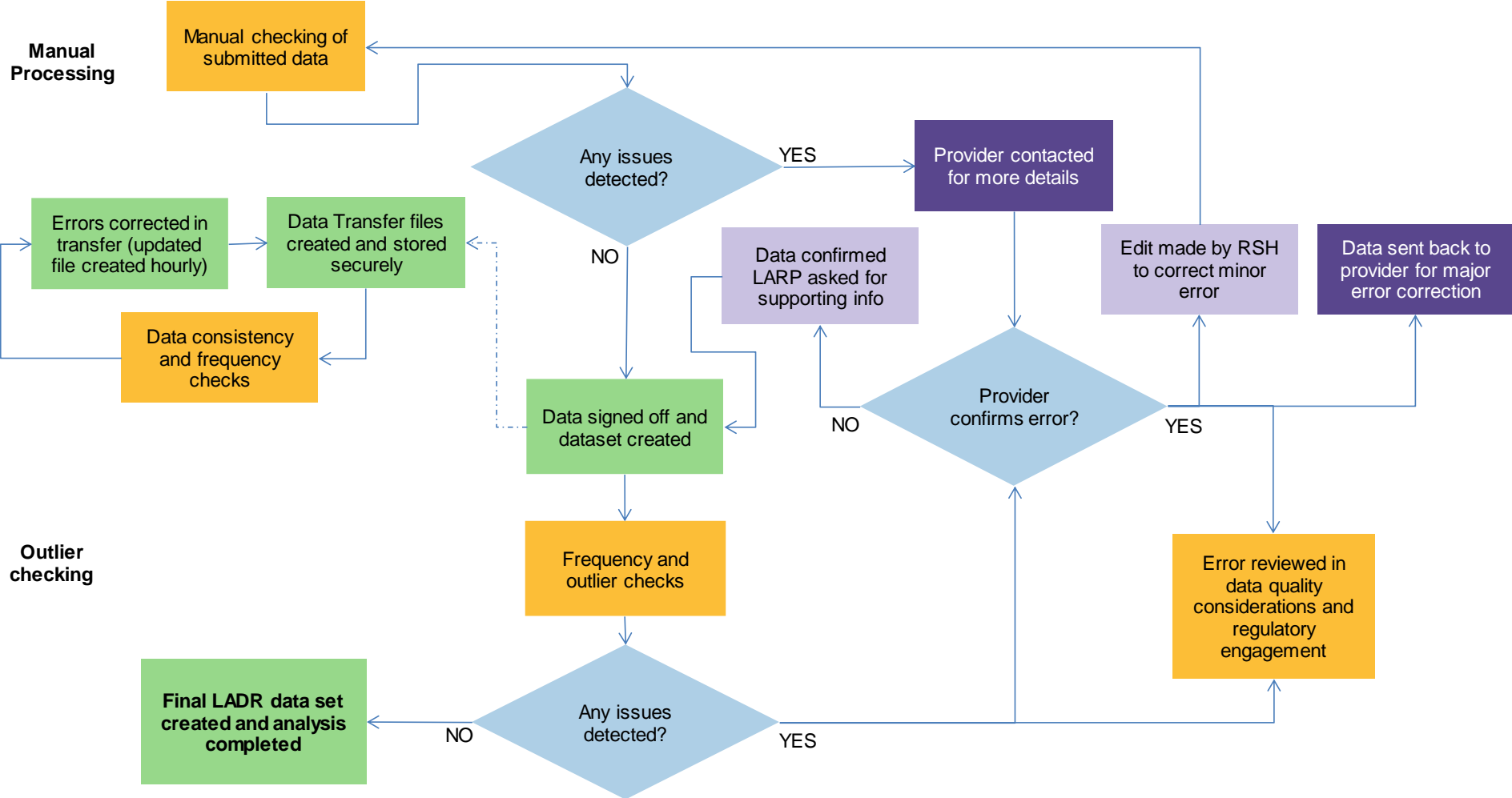**All needs met**          **Some needs met**          **No needs met**

# Annex A: Quality assurance processes – process map

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

**System Processing - Import**

Data Imported by LARP → Basic tests on import integrity and expected data entry → Is data in template valid?

Is data in template valid? — NO → Error message displayed and data does not import

Is data in template valid? — YES → Data saved into database and displayed onscreen

**System Processing - Data Entry**

Data saved into database and displayed onscreen → Does data need editing?

Does data need editing? — NO

Does data need editing? — YES → Provider edits data onscreen

Data entered into system directly by LARP → Input meets cell requirements?

Provider edits data onscreen → Input meets cell requirements?

Input meets cell requirements? — NO → Error message displayed and data not saved

Input meets cell requirements? — YES → Data saved into database and displayed onscreen

**System Processing - Validation**

Validation process activated (individual or all parts) → Data passes validations?

Data passes validations? — NO → Error messages displayed. Provider asked to review → HARD validation error (submission prevented)

Error messages displayed. Provider asked to review → SOFT validation error (review suggested and supporting info requested)

Data passes validations? — YES → Validation process complete data can be submitted → Provider submits the LADR data

OFFICIAL

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

**Manual Processing**

Manual checking of submitted data

Any issues detected? — **YES** → Provider contacted for more details

Errors corrected in transfer (updated file created hourly) → Data Transfer files created and stored securely

Data consistency and frequency checks

Data confirmed LARP asked for supporting info

Edit made by RSH to correct minor error

Data sent back to provider for major error correction

**NO**

Data signed off and dataset created

Provider confirms error? — **NO** / **YES**

**Outlier checking**

Frequency and outlier checks

Error reviewed in data quality considerations and regulatory engagement

Any issues detected? — **NO** → **Final LADR data set created and analysis completed** / **YES**

**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 and maintain homes of appropriate quality that meet a range of needs.**