Data Quality and Methodology

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

2019-2020

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 will use the LADR data extensively as a source of administrative data. We first collected this data in 2020 and will use it to deliver our operational approach to regulating rents (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 are publishing these data.

This is the first publication of these data as collected through the LADR. In previous years, data on owned local authority stock and rents were collected through the local authority housing statistic (LAHS) and published by MHCLG in the corresponding statistical release. For more information see: https://www.gov.uk/government/collections/local-authority-housing-data#2019-to-2020

National Statistics status

The statistics derived from the LADR data and published as local authority 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

In 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 MHCLG to 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, commenting specifically on the:

- Close working with MHCLG both before transfer of collection and on an ongoing basis, to ensure comparability, quality and coherence of the data collected.

- Enhanced quality of the data and improvement in the response rate through building relationships with local authority data suppliers. With data quality being explicitly reinforced through both our in-year regulatory work and the emphasis we place on accurate and timely data, as highlighted in our Regulating the Standards document.
- Enhanced insight provided by our new registered providers' bulletin, which provides the fullest picture of stock and rent information across local authority and private registered providers social housing provision in England.
- Increased granularity in the statistics, with the provision of social and Affordable Rent information split by general needs and supported housing, and tables, data and tools, allowing for further interrogation of the data.
- Renaming of the statistical outputs so that they refer to the subject matter covered rather than the data collection used.
- Transparent and thorough documentation on quality, methods and the strengths and limitations of the data, including the comparison table explaining the nature and extent differences in the new data compared with that previously collected by MHCLG, and a process map setting out RSH's quality assurance approach.

They also identified some areas where the public value of the statistics could be further improved, to continue to maintain the high standards required of National Statistics. We therefore undertake to:

- Publish information about our ongoing approach to user engagement to support further contact with users and future developments to the statistics (within the next 12 months).
- Consider how to better communicate any uncertainty in the statistics, for example by considering confirming the self-reported nature of the data or using rounded numbers (particularly if imputation is needed in future years).
- Work to explain key terms in the statistics when they are first used to enhance accessibility for non-expert users in our next publication (Autumn 2021).

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 the Autumn each year, with the data pre-announced in the release calendar. However, in 2020/21, due to our response to the Coronavirus pandemic, where we extended the survey deadline from 17 July to 31 October 2020, the release of these statistics was delayed until February 2021.

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 MHCLG live tables and our registered provider social housing statistics.

Quality assurance of administrative data (QAAD)

The data used in the production of these statistics are classed as administrative data. In 2015 the UK Statistics Authority (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 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 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 underpins the collection of LADR data. For more information about us and the regulatory framework please see https://www.gov.uk/government/organisations/regulator-of-social-housing.

Regulatory standards

First collected in 2020, the LADR data will be used to facilitate our operational approach to regulating the rent standard from 1 April 2020.

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. In 2020, as part of our response to the Coronavirus pandemic, the deadline for the LADR was extended to 31 October from than the originally published date of 17 July.

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 data set 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 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 LARPs is redacted within the public release to remove all contact information submitted within the Entity Level Information (ELI) 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 us (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. On 1 April 2020 the LADR was included on the single data list which outlines data sets local government must submit to central government https://www.gov.uk/government/publications/single-data-list.

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 statistical data return (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 MHCLG (see https://www.gov.uk/government/collections/local-authority-housing-data) collects 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. As part of this process we worked closely with MHCLG to ensure continued comparability year-on-year. However, there were some necessary amendments to the questions to enable us to undertake effective regulation of rents:

- Local authorities are asked to record details of stock and rents for all social housing they own in LADR, and not just that within a housing revenue account (HRA) as in LAHS. The impact of the inclusion is minimal (less than 1,500 units (0.1% of social stock)) as most social housing units are held within HRAs.

- The data collected in LADR provides an increased granularity in information when compared to that previously collected in LAHS, providing social and Affordable Rent information split by general needs and supported housing. These data have been recombined in the statistics where necessary to provide ongoing trend information.

- LAHS collected 'dwelling equivalent' data for units within hostels and houses of multiple occupation and this is replicated in the LADR return to ensure continuity in the series. However, LADR also collects the number of non-self-contained bedspaces and the average rents charged to provide a more detailed breakdown of this type of provision.

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

Systems

The LADR 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 LARPs.

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 the NROSH+ website 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 the NROSH+ system 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.

In 2020 we also used the 2020 provisional LAHS headline stock data to assess the quality and coverage of the LADR data. We identified a potential 0.3% divergence between total low cost rental stock levels recorded in LADR and LAHS (with LADR recording fewer). We reviewed individual LA discrepancies against queries raised during cleaning to identify reasons for these. Where we did not have information, we contacted LARPs to confirm the data submitted to LADR was accurate. Data from LADR has been provided back to MHCLG 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 the **technical notes and definitions.**

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 local authority registered provider 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.

2020 responses

In 2020 the overall response rate was 100%. This is a slightly higher response rate than for the LAHS return in previous years (98.4% in 2019 and 99.4% in 2018).

Due to the 100% response rate and all data queries being resolved in 2020 we have not needed to impute any figures to account for of 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 is 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 years' release.
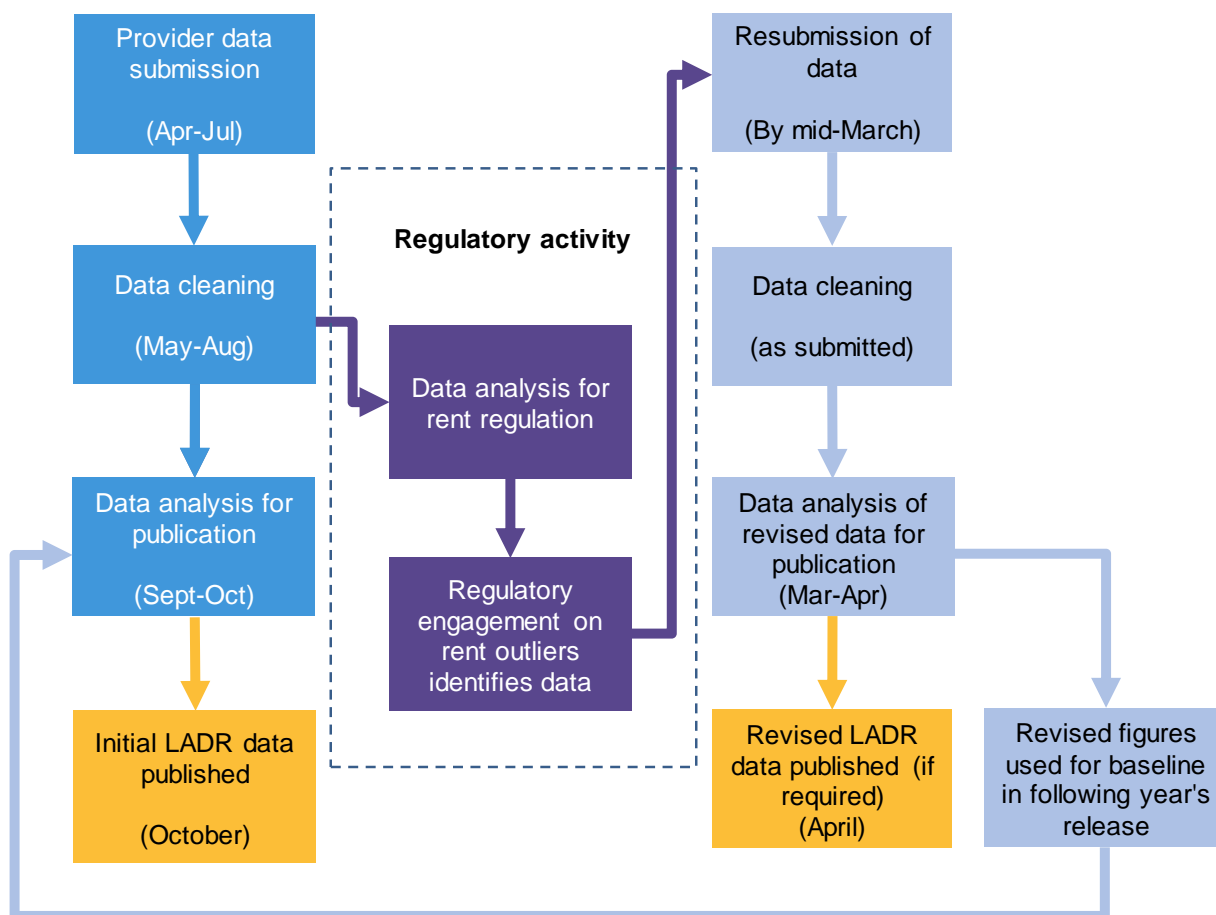
Why not have your say on our statistics in 2021/22?

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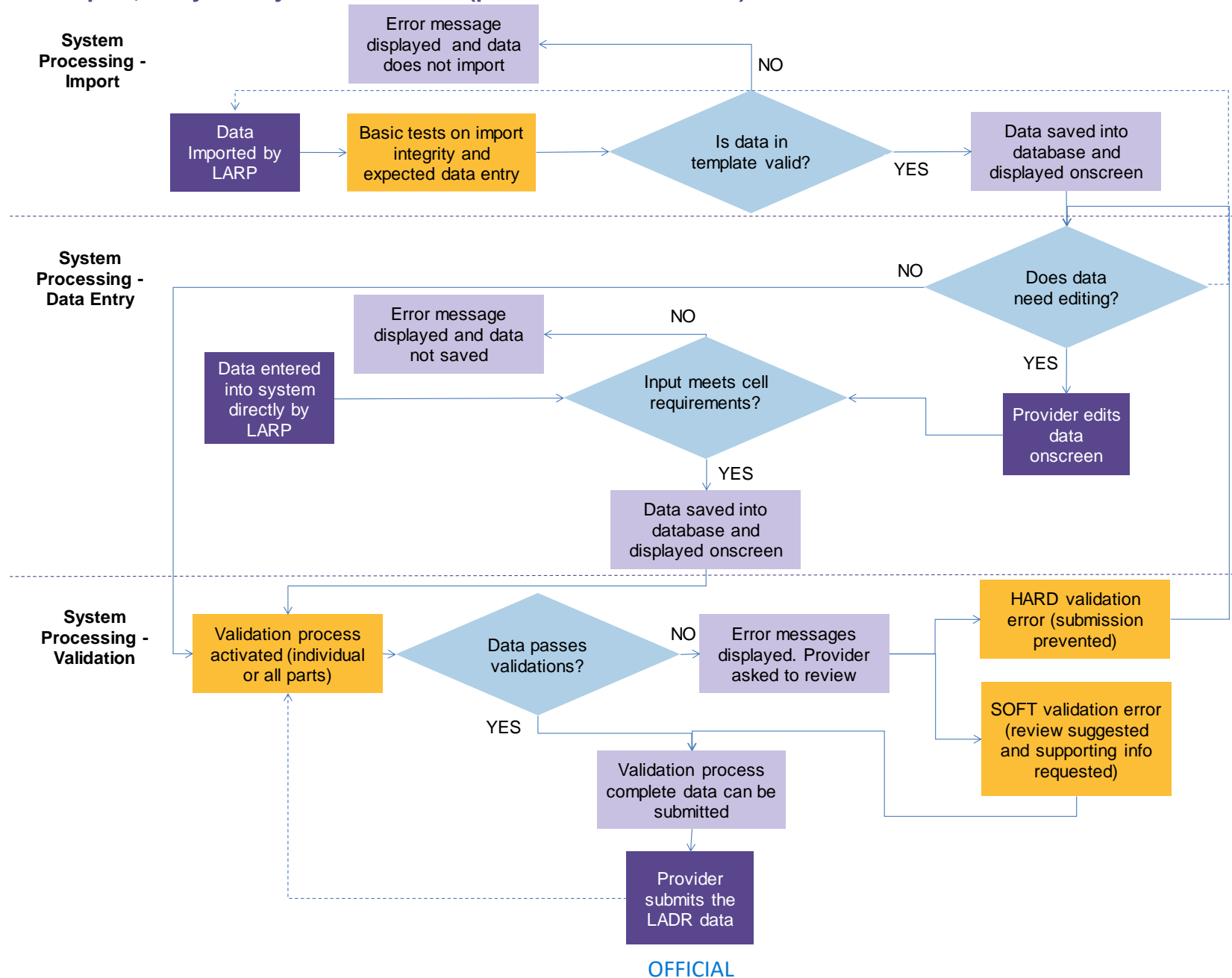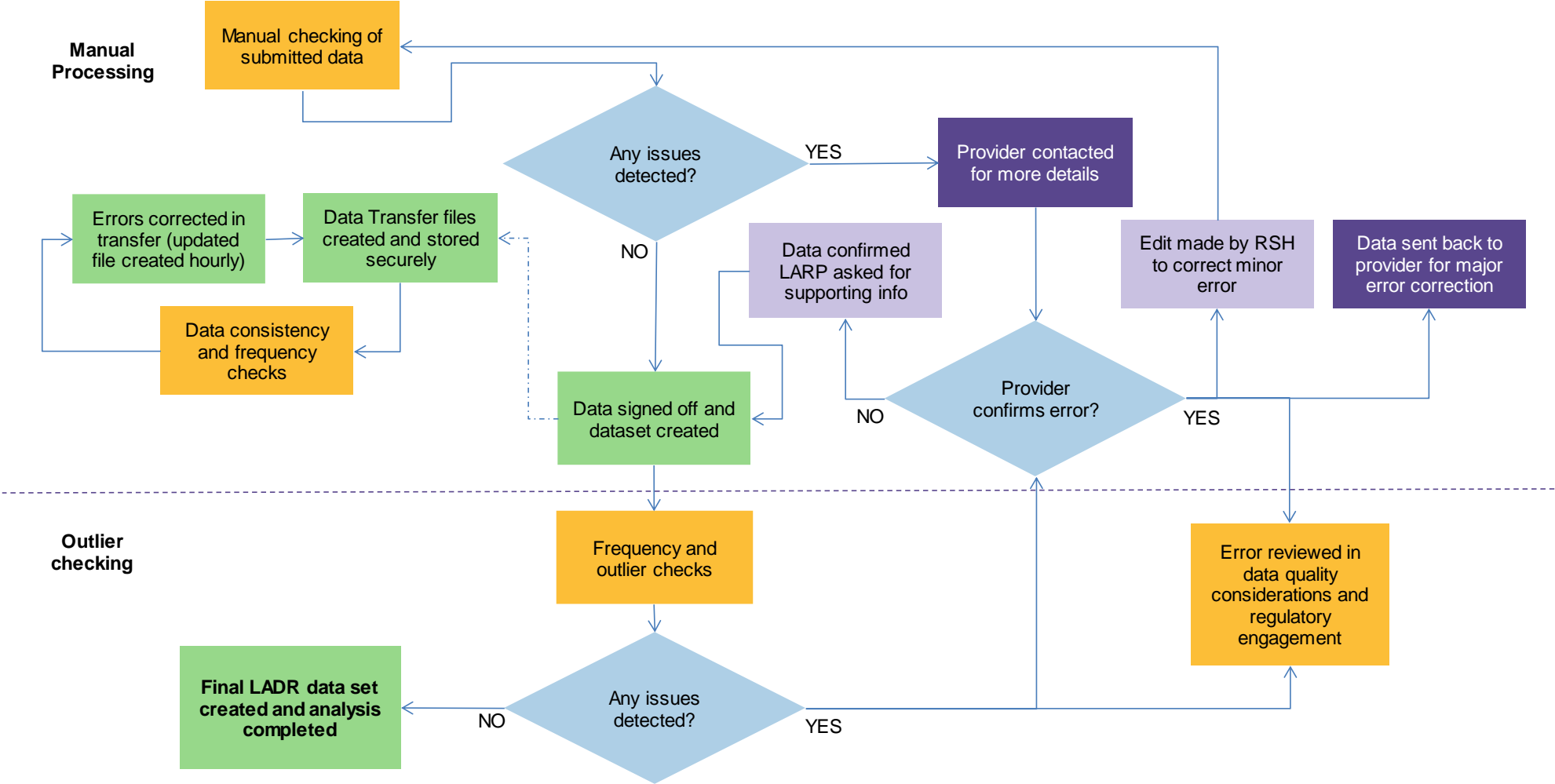
All needs met　　　**Some needs met**　　　**No needs met**

Annex A: Quality assurance processes - process map

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

System Processing - Import

Error message displayed and data does not import

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

NO → (Error message displayed and data does not import)

YES → Data saved into database and displayed onscreen

System Processing - Data Entry

Does data need editing?

NO →

Error message displayed and data not saved

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

NO → (Error message displayed and data not saved)

YES (from Does data need editing?) → Provider edits data onscreen → 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?

NO → Error messages displayed. Provider asked to review → HARD validation error (submission prevented)

SOFT validation error (review suggested and supporting info requested)

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**

NO

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

Data signed off and dataset created

Provider confirms error?

NO

YES

Edit made by RSH to correct minor error

Data sent back to provider for major error correction

Outlier checking

Frequency and outlier checks

Any issues detected? — YES → **Error reviewed in data quality considerations and regulatory engagement**

NO → **Final LADR data set created and analysis completed**

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.