

# PASA Data Working Group

Produced in partnership with:

**isio.**

PASA Experts for Data

**Data Presence v Accuracy**

**June 2024**

# Contents

Section	Content	Page
1	Introduction	1
2	Why is accurate data required?	1
3	What is accurate data?	2
4	What trustees can do to improve data accuracy	2
5	Future proofing data accuracy	5
6	Summary	5

# Acknowledgments

PASA is grateful to the authors of the Guidance and members of the PASA Data Working Group and their employers.

Chris Tagg (Board Sponsor)	<b>PASA Board Director</b>
Kristy Cotton (Chair)	<b>Isio</b>
Paul Rickman	<b>LCP</b>
Steve Robinson	<b>Equiniti</b>
Adam Scarff	<b>Nestle</b>
Alan Clay	<b>LexisNexis Risk</b>
Claire Fuller	<b>First Actuarial</b>
Hannah Blomfield	<b>Intellica</b>
Julie Beadle	<b>WTW</b>
Louise Donohue	<b>Heywood</b>
Nick Jones	<b>PCSL</b>
Paul Young	<b>Capita</b>
Tim Drye	<b>Mortality Manifest</b>

## 1. Introduction

This Guidance provides insight to why data shouldn't only be present, but accurate. It considers what accurate data is, what trustees should do to improve their data accuracy and suggests steps they should take to ensure data accuracy is maintained.

## 2. Why is accurate data required?

Pension schemes navigate through an ever-changing landscape, meaning trustees need to be prepared to face new obligations in addition to existing duties. These requirements necessitate data to be present and accurate. An increased focus on good record keeping is being taken by The Pensions Regulator (TPR), and what was once considered a 'nice to have' is now a 'must have' for schemes, including cost conscientious trustees.

Many aspects of scheme management require complete and accurate data, e.g. dashboards requires an elevated level of data accuracy for key member details ('Find data') utilised by the Identity Service to enable members to access their records. Schemes with inadequate data accuracy will have higher levels of queries for their administrator, increased costs, low levels of engagement or even worse, data breaches. Accurate data underpins the analytics required for reporting and high-quality decision making both for trustees and their advisers for all long-term strategies. Other benefits of accurate data include:

- Risk transfer exercises and wind-ups can be achieved faster, with potentially lower premiums
- Reduced ongoing administration costs as less data checking is required on an ad hoc basis when processing member events
- An improved service with more effective communications and less complaints about data errors
- Increased potential for straight through processing and online transactions
- Facilitation of self-service tools by members and digital engagement
- Reduction in incorrect pension payments caused by poor data, lowering the risk of expensive future remediation
- Lower risk of reputational damage for trustees and employers
- Improved scheme valuation data and funding strategies
- Reduction in ongoing PPF fees or transfer costs

PASA has previously published [Data Accuracy Guidance](#) explaining why individual data items are important.

Aside from the positive benefits to trustees, employers and pension savers, there are also explicit obligations on trustees to ensure the accuracy of pensions scheme records are maintained. TPR requires regular reviews of a scheme's data quality through annual production of common and scheme-specific dates scores, or at key scheme changes such as de-risking or wind-up.

UK GDPR also requires all reasonable steps be taken to ensure personal data is accurate. [Article 5\(1\)\(d\) of the UK GDPR](#) requires personal data to be:

accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy').

To assist managing data effectively, trustees should produce and maintain a [Data Management Plan \(DMP\)](#). A DMP formalises a scheme's key data considerations, the responsibilities of its key stakeholders and documents the policies and controls in place for managing data effectively.

### 3. What is accurate data?

In order to assess data accuracy, the criteria for a data item to be deemed accurate needs to be precisely defined. Many trustees will have been engaged in data quality reviews, but the accuracy of data should be seen as a distinct element of the overall data quality process. This is one of the six data quality dimensions as defined by the [Data Management Association UK \(DAMA UK\)](#).

Data accuracy isn't just a component of data quality, it's the most important characteristic which makes data usable and purposeful. Data accuracy, as the essential standard of data quality, refers to '[the consistency of data with reality](#)'. Data may have been reviewed as part of a common data report and found to exist and be of reasonable quality, but this is different from it being accurate. E.g. schemes may hold a date of leaving pensionable service, which is present and within the standard data conventions, but isn't the accurate date the member left the scheme. User input errors, poor data controls or reformatting because of system changes can result in inaccurate data changes which remain undetected by standard data completeness checks. Another example would be a member having a date incorrectly transposed where the date of leaving pensionable service is 12/01/2022 but is input as 01/12/2022. The data passes system validation checks but results in the member being in receipt of 11 months less benefits than they're entitled to at retirement. If this was reversed the trustees would fund 11 months' of benefits the member isn't entitled to.

### 4. What Trustees can do to improve data accuracy?

Improving data accuracy can only be achieved once the issues in the data are known, and the root causes of those issues are identified. Trustees should undertake a data quality audit to identify issues in their data and areas of weakness, and review and remediate issues where necessary. One of the most common issues facing schemes is missing data. Where precise data can't be located, e.g. date joined scheme, it's important this is clearly documented. If the system is updated with an assumed date, there must be a decision process and clear audit trail. Consideration should be given to the advantages, disadvantages, and consequences of using

---

<sup>1</sup> Article, 'Why is data accuracy important' by Omrean - <https://omrean.com/why-is-data-accuracy-important/>

assumed data. In some instances, a blank data field where data is missing is an acceptable, and sometimes more likely to flag an issue than using assumed dates.

Determining what accuracy analysis can be performed is the first step. Schemes can test data for validation criteria which can help to identify where data held could be incorrect. For example, testing for 'dummy dates' or 'filler data'.

Tracing of member identification data is important and can be done with help [of third-party agencies](#). It will then be important to consider the data held, the source of the data and how accuracy can be tested. A key approach is analysing the consistency of data between different data sources. In most cases these will be limited to digitised records and can be analysed by data analysis tools.

Ensuring data fields are cross-referenced to other related fields enables validation which will identify data inaccuracies. Scheme salaries, contributions and other service or transferred in data can all provide context to provide greater assurance a single date joined scheme isn't only present and logical, but also accurate. Therefore accuracy needs to be viewed in context of the wider level of data accuracy.

Lack of cross-referencing to non-digitised data can present an issue for older records which are more likely to have suffered data degradation over the years through manual input and system migration. A review of all data stores should be undertaken to maximise the level of successful data analysis. This may involve the digitisation of paper files and microfiche, but also any additional data held outside of core member records, such as Excel files or other computerised records. In all cases, the data should be stored on a single electronic administration platform to maximise the effectiveness of the data accuracy checks and to ensure data is stored in a safe and secure manner.

The tools used for data quality analysis are varied; from internally developed spreadsheet-based data reviews or those built into the administration systems themselves, through to bespoke boutique suppliers offering data analysis services with system or even client specific customisation. Each of these systems will need to be developed to cover the bespoke nature of each scheme and have their own set of data quality rules developed alongside a review of the Trust Deed and Rules. Historic corporate activity will also need to be considered as this may affect the structure of the data held for different subsets of members in a scheme when compared to the core of the scheme membership e.g. historic bulk transfer arrangements or one-off benefit augmentations. Each scheme needs to be reviewed on its own merits and requirements.

There will be two key outputs of analysis flowing from a review of the data accuracy; a trustee report outlining the key issues plus a list of members with these issues identified. The trustee's report should cover a number of key areas:

- **Scale or benchmark of the data accuracy**

This should show an overall assessment, benchmarking against a basket of similar schemes or the scheme's historic data accuracy scores over time. It's important to be able to contextualise the size of any issues

- **Results**

These should be summarised, and the report should show an overall score, along with results for individual data. Using a scoring system such as red/amber/green or percentage ratings for each test run can be helpful. In addition, tests can be prioritised for resolution to allow trustees to make immediate decisions, for example by splitting tests into Critical, High, Medium or Low.

- **Categories of concern**

These should be highlighted in the context of their impact on members' experience and trustee objectives

- **Impact of the data accuracy on the running, funding or derisking of the scheme**

This should consider the specific data sets affected and exclude any members who don't require any remediation work to avoid wasting effort

- **Recommendations and prioritisation of actions**

This may cover urgent intermediate steps which can't wait until a full data remediation project is in place, e.g. prioritisation of data for members coming up to retirement where incorrect benefits could knowingly be put into payment

- **Timescales required for short, medium and long term actions**

This should consider more immediate requirements such as valuations, benefit statements, pension increases or other data project work. Sampling may be appropriate to understand the potential success of a program of work if it's within a wider project

- **Cost benefit**

To include a pragmatic consideration of the cost of the remediation work against the requirements and criticality of the data issues identified and where alternative solutions can be applied

Once the trustees have accepted the recommendations, remediation should be undertaken. To facilitate this, the requirements and solutions will need to be agreed, considering the recommendations of the report and aligning the correct resources and tools.

In addition to data remedy, benefit rectification may be required as a consequence of correcting data or identifying missing data.

Data remediation projects should consider how inaccurate data has arisen and whether any additional controls can be put in place to mitigate future issues. Some may be historical issues associated with system changes or administration provision being moved. Some may be errors not captured at the input stage such as manual input validation which could be improved by reviewing where training or strengthening of system controls can be improved. There may also be bulk processes such as monthly contribution or annual salary data loads and data interfaces from external parties where errors are imported into the administration system if insufficient controls and validations don't form part of an overall data governance framework covering data accuracy.

## 5. Future proofing data accuracy

Steps should be taken to ensure data accuracy can be monitored and where necessary, improved on an ongoing basis. It shouldn't be treated as a one-off exercise. Consideration should be given to:

- Repeatable data validation tests
- Refining tests on a periodic basis
- Avoiding duplication
- Maintaining a DMP
- Maintaining an improvements log
- Maintaining a clear audit trail of when data was last updated and tested, in addition to the relevant decisions made

PASA's [Data Controls Guidance](#), provides further detail on this.

## 6. Summary

The timing of a data accuracy review will be influenced by decisions made by the trustees on the future of the scheme. In the case of de-risking or review of scheme funding, this will be a key driver in the decision making. Data accuracy should be a constant and continual process in the journey to deliver the benefits and service to scheme members.

Trustees should recognise data quality management is an ongoing process and not a one-off exercise. Data needs to be consistently monitored throughout the life of the scheme and controls reviewed to ensure they remain adequate to prevent future data degradation.

Data remediation projects are a useful tool for maintaining accurate data, but trustees should also maintain data improvement and management plans to ensure continuous monitoring. Poor quality data can lead to poor decision making and needs to be given the importance it deserves as the lessons to be learned from not doing so can impact employers, trustees and members for years to come.





Get in touch:

[info@pasa-uk.com](mailto:info@pasa-uk.com)

[www.pasa-uk.com](http://www.pasa-uk.com)