PASA
assess | aspire | achieve

Produced in partnership with:

Deloitte.

PASA Experts for Data

PASA Data Working Group

Ongoing Data Management & Controls Guidance

November 2021

Data Management & Controls Guidance

Acknowledgments

1. Introduction

PASA released Guidance on **Data Management Plans** in March 2021. A Data Management Plan formalises the key data considerations for a pension scheme and the responsibilities of its key stakeholders, and documents the policies and controls in place for managing data effectively. This document builds on the Data Management Plans Guidance and considers the ongoing data management and controls of pension scheme records.

Data management of pension scheme records is an ongoing process which should be considered an activity in its own right. The accuracy of pension scheme records is fundamental and the foundation for the majority of activities a scheme needs to undertake throughout its lifetime.

Pension scheme data only remains accurate for a finite period and will become obsolete if sufficient controls aren't in place to maintain it. A database needs constant management; for example when additional member contributions or pensionable earnings are received for active members, when people leave employment or move home.

If pension scheme data isn't accurate and kept up to date, time, resource and budget will be needed to accommodate and fix inaccuracies. Trustees are reliant on third parties and members to keep their data up to date. Pension schemes regularly hold millions of data items and it's essential appropriate processes and procedures are in place to manage this to ensure data remains in good shape.

Accurate data is essential for regular and one-off project activities, and the importance of holding precise data has never been more important. Trustees are facing ever increasing demands for accurate data such as the dashboards' requirements, regulatory reporting, de-risking activities, automated calculations, and self-service. As such, having confidence in an accurate and reliable dataset should be one of the key objectives for all trustees and their stakeholders.

2. Considerations when setting data controls

One of the key components of database management is to ensure data controls are in place. There are a number of steps which need to be taken to enable a successful outcome:

- Impact assessment - to focus effort and fees in the most critical areas (as per section 4 of the 2019 **PASA Data Guidance**)
- Be specific - understand and agree data items you want to test
- Know your datasets - understand where key data items are stored. Is data held on the administration platform, payroll system or other electronic data sources?
- Paper based or microfiche - consider if any critical data is held on paper files or microfiche and decide if this can be converted into an electronic source. Both fiche and paper will degrade over time. Can this be automated or is manual assessment/input required?

- Database - create and maintain one source of truth. Avoid duplication
- Sampling - review example cases to validate issues before updates are made in bulk
- Benefit Matrix - understand what members and benefit categories should form part of the sampling
- Documentation – ensure detailed documentation is maintained for all stakeholders (e.g. Scheme Actuary and administration team). Define each database item and any exceptions to the rule, along with any assumptions used
- Road map - introduce testing in a phased approach. Agree the critical data items to test initially and add more into the process along the journey. While you want to aim to capture every requirement and data item upfront, it's unlikely you'll identify every test you want to perform first time round. Look to continually enhance your ongoing data testing schedule

3. Consideration for new data received

A scheme can receive new data through various sources and channels, with data typically coming from scheme sponsors, such as a pensionable salary and pension scheme contributions, or direct from a member, such as updates to contact details.

It's important to set acceptance criteria for new data items before they're loaded into any database. If a scheme can ensure new data meets a robust set of criteria before it enters its system, it'll have built in safeguards meaning new data items can only improve the dataset and are consistent with other data items held.

Validation checks should be used to assess the precision of new data and highlight areas for further consideration. Data validation and resolution is a key control for receipt of data and should be performed for all new data received. The validation checks performed, and the parameters for issues and escalation will be scheme specific and should be reviewed periodically to ensure adequate data validation exists. Data items may need to be referred back to source or updates made to other data items to resolve inconsistencies. Total control and balance checks should be run to ensure all data has been loaded.

If repeated issues are identified it's advisable to review the data source, e.g. is there a problem with the HR database? Are you best focusing effort to support correct data at source instead of, or in addition to fixing the scheme's dataset?

With online solutions increasingly available to members it's important to ensure consideration is given to which items a member can and can't update. Active members may want to update information, but this could differ from the HR interface data, for example surname or National Insurance Number. The systems could then become unaligned.

Robust processes should be in place to ensure data fails are identified, recorded and discrepancies are dealt with. If investigations are required there must be measures in place to keep track of the steps taken to date and the next action in the process.

Where new data is arriving on a regular basis, this should be evaluated before it's loaded into any live system, with swift action taken to address areas requiring further consideration.

4. Consideration for existing data items

To improve and maintain the accuracy of your existing data items, data quality assessments can be completed on a regular basis. As part of the data quality assessment it's likely validation checks will be required from different viewpoints. For example, by testing the same data item against other data items where there's a dependency or relationship. This could be across systems or tables within a database. These assessments are typically completed by data quality software.

Data quality should be tested frequently, particularly as it's a regulatory requirement to produce your data quality scores on an annual basis for the Pension Regulator. As all pension activities are impacted by data quality it's good practice to undertake data analysis before these activities, such as prior to producing annual benefit statements, applying annual pension increases or calculating annual allowances. Where data quality issues are identified it's essential you implement a Data Management Plan outlining the steps to data improvement.

The data quality tests should highlight where further attention is needed. Consideration is required to ensure effort and budget are initially focused on areas needing the most attention and which will yield the greatest cost benefit to the scheme. Priority should be given to those data items which can be ranked the highest of importance, or areas where there are lots of fails to understand the root cause of the issue.

5. Importance of benchmarking

Benchmarking allows a baseline to be drawn which is a valuable reference point for future data quality assessments. The benchmarking will allow progress to be monitored and improvements to be measured from one period to the next.

Benchmarking software can be aligned to a database to create bespoke tests which take scheme design into account. These tests can then be automated and repeated. Utilising software enables repeatable and cost effective data analysis, enabling stakeholders to monitor data quality and the impact of data improvement techniques. It's common for some degree of tailoring to be required for any benchmarking software to be effective.

As part of this analysis it's best practice to grade the results. For example how many members pass/fail every test, or how many members pass/fail the tests you deem critical. This will allow you to prioritise certain members or tests where there are multiple fails.

> **Trustee and stakeholder reporting**
>
> The administrator should provide regular reports setting out data activity in the reporting period, any areas where improvement work has taken place, areas of concern and a plan for what's being targeted for the next reporting period

> **KPIs and data scoring**
>
> Much like ensuring data tests are consistent, reporting of key performance indicators (KPIs) and data scores should also be consistent from one reporting period to the next. Where tests have been amended this should be highlighted as this will help monitor progress and improvements. It'll also allow stakeholders to decide where to focus effort and spend in the next period. Data scores, particularly common data, may pass a test simply by having something in a data field. This means while the data scores may seem good and give comfort to many, they can sometimes be misleading. Therefore considering 'how' you score is vital.

6. What does regular testing look like?

Regular testing should be:

- Consistent - the same tests are run from one report to the next, with new tests being marked appropriately so they can be baselined the first time round. If tests are amended, decide if a new baseline result is needed
- Measurable - Pass or Fail. Where a fail result requires action. Previous results should be used to monitor progress
- Repeatable - tests can be re-run and aren't a one off exercise
- Frequent - at regular intervals - at least annually, but consider monthly, quarterly or in-line with software updates
- Achievable - set realistic targets and ensure sufficient resource and budget is available

Ongoing data testing may initially be driven by a focus on specific activities such as dashboards or GMP Equalisation readiness. Individual activities will likely be run under a project structure with milestones and an end point in the lifecycle. While individual projects serve a purpose, the aim should be to have these as sub projects forming a wider framework and data strategy encompassing all data related activities you might want to achieve. The ongoing regular maintenance of data then becomes a BAU function in its own right and is managed through the Data Management Plan.

7. Hints and Tips

This section provides some practical hints and tips.

Data collection - targeted versus wide net approach

In any data cleanse exercise, it's likely results, at least initially, will show you'll need to update multiple data items. One consideration will be whether you collect data items at the same time for each member or pick off one data item across all members. Consider running a proof of concept phase to test alternative approaches on a sample of members to see what's the most efficient and accurate way of collecting data. The quickest way may not always be to collect the data in one sitting, especially if the data is spread across multiple data sources. Data gathering isn't just about collection, it often requires an assessment of the member's record and surrounding information. Some data items may need careful assessment before they're updated or added into the scheme's database.

Data exceptions and assumptions

It's likely a scheme will have some data items which don't fit the rules set for the database and database fields. For example, members receiving discretionary benefits, or a transfer-in granted on special terms. Database markers can be added to the database so users of the data, now or in the future, are aware of exceptions and these database items aren't overridden or marked as inaccurate.

If there are missing specific data items or the database has strict requirements for certain fields (e.g. it must always be populated), you may have to make some assumptions or use temporary data items, such as a temporary national insurance number or a dummy date. Consider adding a marker to the database to ensure this is clear to all users and ensure central documentation is updated with this information. If attempts have been made to cleanse certain data items, ensure you have a mechanism to identify when database fields were last assessed.

Quality Assurance

Some data cleanse activities will be automated and others will require manual intervention. Updating data items introduces risks. To reduce the risk of updating the database with inaccurate data or overriding accurate data, ensure sufficient quality assurance checks are in place before the updates are made. Ensure a backup is made before data cleanse work is undertaken and testing happens outside of the live system. This provides the ability to back out any changes, especially bulk changes, before they go live. Recording who, when and why the data is being updated also provides a robust audit trail. A peer review system should operate for manual updates. For example, by sampling the work carried out. Where automated activities are run, ensure these are returning accurate results by sampling the dataset.

Operational capacity

Where activities are conducted to identify and correct bad data, it's vital to ensure there's operational capacity to correct data in a timely manner.

Other considerations

- **Duplication** - Some databases or systems may allow or require data items to be stored in multiple places. Ensure tests take into account all relevant database fields

- **Remarks** – Comments are often used to record data items. It can be worth spending the time to review these in bulk to ensure any key data items within them are held within a pre-defined field in the database rather than as part of a string of text. Remarks can be helpful in certain instances. But it's often more effective to use defined database fields

- **Update to individual data items** - Ensuring updates are made within a well-defined set of procedures is important. However, having a rigid structure in place which, for example, only allows one person, to make updates can be counter-productive. If sufficient controls and checks are in place, ensuring all stakeholders have the ability to update or at least contribute to maintaining a clean data set, provides a better chance of keeping it up to date. Use a log to record data items which need to be updated or reviewed. Ensure SLAs are in place so updates are assessed and completed in a timely manner

- **Member engagement and communication** - It's often beneficial to ensure members know data cleanse work is taking place. This could be explained in newsletters or on a scheme's website. It's good practice to remind members to verify their personal details on a regular basis. This can plug gaps in the dataset, particularly where they can provide proof or the information can be validated against other data items held in the scheme database

- **Visibility** - Providing members with visibility of the data held can be helpful in ensuring it remains accurate. Education and supplementary information may be required to support members' understanding of pension terminology and practices. General Data Protection Regulation requirements allow members access on request to the data a scheme holds. Consider if you can go further with allowing members to see the data which relates to them

- **Reconciliation** – Ensure total and balance checks are completed for monetary values, membership counts and dates. These checks should highlight missing records or values, duplicate records, incorrect values or incorrectly formatted values

8. Conclusion

Data never stands still and accurate datasets underpin the successful management and tasks carried out to meet trustees' duties. Careful consideration should be given when setting data controls, trustees should ensure they have visibility and a full understanding of how data enters the scheme database, is kept updated and amended. Benchmarking of datasets will allow future assessment to happen consistently and will form part of a regular testing schedule.

A data management plan is a management tool to document the controls required for maintaining accurate scheme records. Managing the quality of scheme data is a key trustee responsibility and developing, implementing and maintaining a Data Management Plan which meets the need of the scheme should be an action for all. Focus on data is greater than ever and the pensions industry must take steps to prioritise keeping it accurate.

Get in touch:

info@pasa-uk.com

www.pasa-uk.com