

Ensuring Data Integrity: A Three Step Approach to Clinical Dashboard Validation

River Meckstroth, MMS Holdings, Michigan, US
Kris Wenzel, MMS Holdings, Michigan, US

ABSTRACT

In an era where clinical decisions increasingly rely on automated visualizations, ensuring the accuracy of data within PowerBI dashboards is crucial. This presentation explores an innovative three-step validation process that addresses the challenges of data ingestion, model integrity, and visualization accuracy in clinical dashboards. By employing rival programming techniques and rigorous quality control checks, we mitigate risks of data misrepresentation and uphold the ethical use of clinical data. The automation of this process will improve clinical decision-making timeframes and reinforce the reliability of the decisions. Attendees will gain valuable insights into a systematic approach that not only safeguards data integrity but also sets the stage for industry-wide best practices in dashboard validation.

Introduction

In an era of data-driven healthcare, clinical dashboards have emerged as indispensable tools for utilizing raw data to provide actionable insights. PowerBI dashboards, with their ability to ingest, model, and visualize vast datasets, have revolutionized the efficiency and depth of clinical data analysis. The growing reliance on these automated visualization tools has led to a critical need for rigorous validation to ensure their accuracy and reliability. Without proper validation, dashboards risk propagating errors that could undermine clinical decision-making and limit meaningful insight.

The significance of PowerBI dashboards lies within their ability to create dynamic and interactive visualizations that can reveal trends, patterns, and anomalies in real-time. Their speed and flexibility enable healthcare professionals to make prompt, data-informed decisions. It is this level of flexibility that requires extreme thoroughness when validating dashboards. Validation must be robust yet efficient while upholding the same standards as traditional clinical reporting tools to accommodate the active and ever-evolving nature of PowerBI applications.

Currently, the lack of established industry standards for dashboard validation presents a substantial challenge. Validation processes often rely on manual, time-intensive methods that vary widely across organizations, leading to inconsistencies and vulnerabilities. Common pitfalls include inadequate checks on data ingestion, oversight in model relationships, and insufficient testing of visualization transformations. These gaps not only hinder the reliability of dashboards but also raise ethical concerns regarding the responsible use of clinical data. To address these challenges, this paper introduces an innovative three-step validation framework designed to instill trust in PowerBI dashboards. By employing data ingestion validation, model integrity verification, and visualization accuracy checks that are each tailored to mitigate risks at critical junctures of dashboard development, this framework aims to uphold data integrity, enhance clinical decision-making, and set a foundation for industry-wide best practices in dashboard validation.

Validation at MMS

Since there are no established industry standards for dashboard validation, much of the intensive process is left up to self-determination. There are different stages of dashboard construction, and each introduces possibilities for potential inconsistencies to arise. While there are concise data standards that must be upheld, dynamic data visualizations are trickier and require a more comprehensive approach. Combining these different stages into a single process is difficult, but very necessary. As a result, MMS approaches validation through a linear three-step process to limit inconsistencies and errors. The first of which is data ingestion, followed by model integrity, and then visualization.

Data ingestion involves checking the data that is loaded into PowerBI is not incorrectly altered. Model integrity requires ensuring that the binding connections between data tables will not cause potential issues. Visualization validation is required to ensure that the interactive elements of a dashboard are accurately representing the clinical data they are affecting. *Figure1* demonstrates validation moving left to right. If the process uncovers issues with data tables, then there is no reason to continue to validate for model integrity or visualizations. Thus, the validation process should move linearly.

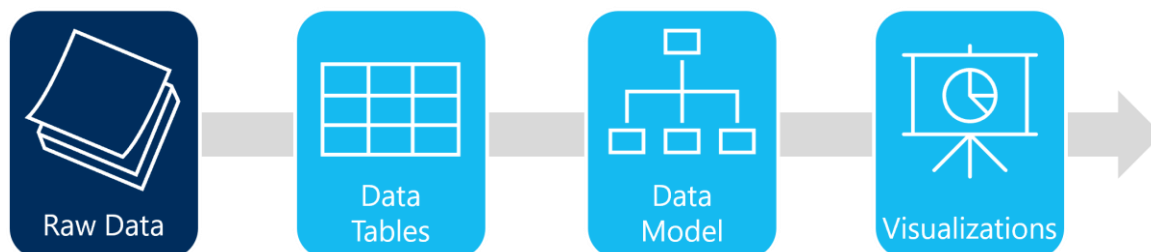


Figure 1 – Linear Steps of Validation

Dashboard Visualization Setup

Clinical dashboards are built using a combination of structured data models that are used to build interactive visualizations. Ensuring the integrity of these dashboards requires a design approach that prioritizes accuracy, usability, and scalability. An effective clinical dashboard should attempt to maintain clarity and reliability by utilizing simple and straight forward operations and principles whenever possible. A key principle is keeping computations within the dataset rather than relying on complex real-time calculations in PowerBI. By handling derivations and relationships at the dataset level whenever possible, consistency is maintained across reports and there is reduced possibility for errors or oversight to occur. Additionally, leveraging statistical functions that have been validated by Microsoft ensures that basic metrics and transformations adhere to industry standards. From a development standpoint, there are many foundational rules that can limit potential issues. While following these rules can limit the amount of validation necessary, vigilance is important when it comes to the consistency of the validation process.

A well-structured dashboard consists of several key elements, each contributing to data integrity:

- **Underlying Relationship Model:** Defines the bindings between datasets to ensure accurate connections, filtering and aggregation.
- **Application of Data Visualizations:** Includes charts, tables, and other graphical representations to highlight critical insights.
- **Filters and Features:** Enables users to dynamically interact with data through slicers and drill-through functionalities.
- **Individual Visualizations:** Each dashboard is composed of multiple visual components that must be validated independently.

Dashboards must be adaptable to meet different clinical needs while maintaining standardization where possible. Standardized dashboards based on common data models improve efficiency and comparability across studies. However, variations in data sources—such as raw Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) datasets—necessitate custom dashboards that cater to specific research requirements. Ensuring validation across these diverse setups is crucial for maintaining accuracy and reliability. By adhering to these principles and components, clinical dashboards can provide meaningful insights while ensuring data integrity and usability across different healthcare applications.

Assessing Dashboard Quality

The effectiveness of a clinical dashboard is not solely determined by its ability to visualize data, but also by the accuracy, usability, and reliability of its insights. Ensuring dashboard quality is critical to preventing data misinterpretation and supporting clinical decision-making with confidence.

A high-quality dashboard must excel in several areas, including but not limited to accuracy, usability, responsiveness, and aesthetic considerations. It is important that the displayed data precisely reflects the underlying dataset without distortions or transformation errors. Furthermore, dashboards should be usable and intuitive, allowing user to navigate and extract insights without extensive background. A dashboard's responsiveness when handling large or complex datasets is essential, performance should be able to remain optimal despite these challenges. Clarity and proper aesthetics are critical to avoid confusion or misinterpretation of data. Ensuing dashboards are upheld to these criteria ensures that clinical professionals can confidently rely on the insights they provide.

Furthermore, dashboard effectiveness is a very important measure. To ensure dashboard effectiveness, standardized evaluation methods are employed, including data integrity checks, which compare dashboard outputs against validated raw datasets to build confidence that any transformations have been correctly applied. Additionally, performance benchmarks assess load times and filter responsiveness to establish usability in real-world applications. Validation reports involve implementing automated validation logs to provide transparency and traceability for potential issues. By leveraging these assessment techniques, we ensure that dashboards meet the rigorous standards required for clinical use.

As dashboard validation evolves, opportunities to enhance quality assurance arise. Continuous monitoring of dashboard functionality, regular validation updates, and standardized metrics will contribute to maintaining high-quality standards. Addressing validation gaps and integrating industry-wide best practices will further strengthen the reliability of clinical dashboards.

Ensuring dashboard quality is a fundamental step in fostering trust in clinical data analytics. By prioritizing accuracy, usability, and efficiency, we can create dashboards that serve as reliable tools for informed decision-making in healthcare settings.

Identify Validation Opportunities

It is important for dashboards to be built with regards to who will be using them, using content and formats that will be most likely to provide quick and useful insight. According to National Clinical Audits, dashboards are designed to "provide feedback to clinical teams and managers, in order to monitor care quality and stimulate quality improvement. However, there is limited evidence regarding the impact of quality dashboards and audit and feedback research focuses on feedback to individual clinicians, rather than to clinical and managerial teams" (reference 1). While dashboards can severely impact decision making for managerial teams, often their use is best served for individual use. Recognizing this disparity allow a shift to designing dashboards that will be useful for clinical teams, as impacting their decision-making has led to more decisive action.

Building and validating dashboards while focusing on the correct audience will increase reliability and strength of clinical dashboards. Effective validation of clinical dashboards requires recognizing areas where quality assurance can be strengthened. By identifying validation opportunities, we can proactively address potential issues before they impact decision-making.

Validation Methodologies and Approach

The importance of being thorough cannot be overstated when it comes to clinical validation. As previously established, there are a lack of validation standards in the pharmaceutical industry. Thus, there is a clear necessity for strong and established principles that can be used as a road map going forward. A thoughtful and thorough approach to validation becomes much more important when it will be used to define these previously untouched standards.

The MMS validation approach is designed to cover all the elements of dashboard creation and use that could be incorrectly altered or skewed by dashboard building software. As demonstrated above in *Figure 1*, the validation process will flow linearly through the steps of dashboard creation, beginning with data ingestion, moving to model integrity, and finishing with visualizations. Further explanation about the steps of the validation process will be discussed in detail below.

Because the thorough approach to validation currently requires intensive work that must be repeated for each dashboard, MMS is attempting to streamline and renovate the process using automated tools. Automating some of

these tasks would save countless hours of repetitive work. However, with the cutting-edge software and tools available, automation has now become possible.

Rival programming is the primary technique used to validate dashboards and can be utilized extensively when it comes to automating some of the more intensive processes. A rival program is designed to replicate the work that is done behind the scenes when using dashboard design software like PowerBI. By using simple programming tools that can be trusted, programmers retrieve results that should match those generated by PowerBI services. Any differences in these results could be seen as oversight by PowerBI.

Python's vast data manipulation tools make it the frontrunner for most rival algorithms. Tools such as the PowerBI query editor and visualization builder can be replicated by some of python's data-based modules. When utilized correctly, these tools can be used for automation that could revolutionize dashboard validation.

The Validation Process

The validation framework follows a structured, stepwise methodology designed to ensure accuracy at every stage of the dashboard lifecycle. Our approach focuses on the three steps of dashboard creation data ingestion, model integrity, and visualization building.

Data Ingestion

Data ingestion occurs when PowerBI loads in the data that is used to build the clinical dashboard. During this process, a user can alter the data by using the power query editor. The query editor has a variety of functions and can change a data table drastically.

However, it is necessary to validate against the abilities of the editor to ensure that there is no oversight during these large transformations. Accidental dropping of rows, improper merges, and misrepresentation of clinical data are all examples of possible mistakes the PowerBI query editor can make, and validation is required to catch these errors when they occur.

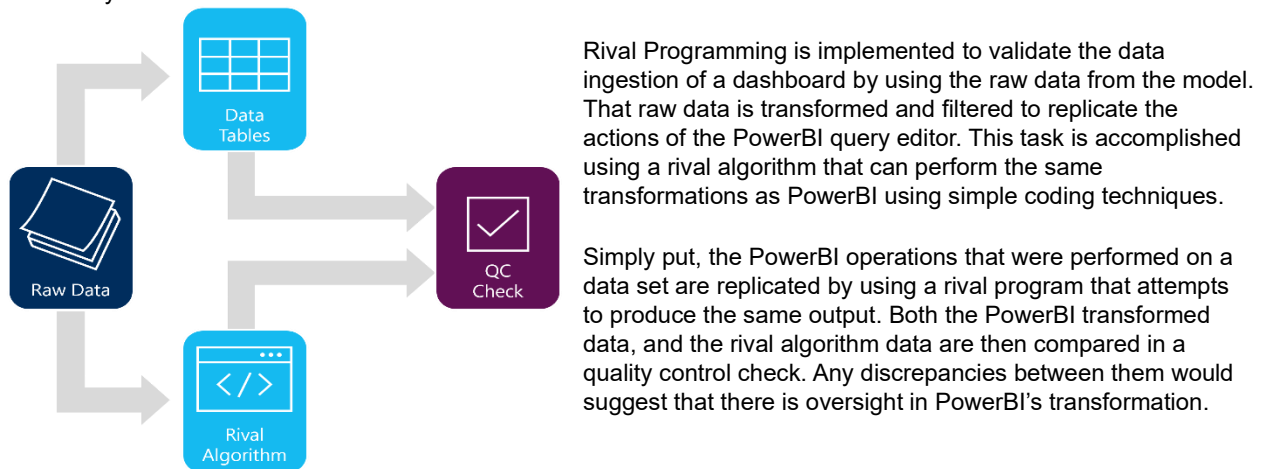


Figure 2 – Validation of Dashboard Data Ingestion

Model Integrity

When there are multiple tables within a PowerBI report, the connections established between these tables create a model. The PowerBI model view allows for visualization of all the different properties and connections within a dashboard. Validation of a dashboard's model integrity does not require any form of testing or rival programming. Instead, a visual inspection is performed. There is an individual inspection on each relationship, the purpose of which is to check that the proper columns are linked together. Furthermore, there may be red flags, such as the "many-to-many" cardinality, that cause persistent issues within the dashboard. If these red flags pop up, it is the established practice to check with the designer of the dashboard that they are the intentional.

Visualizations

The last stage of validation is undoubtedly the largest and most time-consuming. PowerBI visualizations are built to represent data in many ways. They are built using columns from the ingested data. The columns are transformed and represented as different types of charts, tables, or graphs. After they are built, visuals can be temporarily altered by different slicer and filter tools depending on the users' desires. These tools are extremely powerful and frequently used; thus, it is necessary to ensure that these methods are accurately representing the data they are affecting. Validation is required to ensure that these PowerBI transformation tools do not misrepresent the visualized data and lead to incorrect assumptions.

A standard PowerBI visualization has so many ways that it can be filtered, and to test against ever different state of the visual would be unreasonable. Thus, visualizations are validated using a configuration file. The config file contains a small sample of different ways to filter visuals, otherwise known as slicer test cases. Each individual test case has several slicer states that are used to filter a page of the dashboard, impacting every visual on the page. The size of the config file and number of test cases varies based on the size of the dashboard and a user's preferences.

A rival algorithm is used to replicate the filtering operations and build filtered datasets identical to PowerBI visuals using the raw data. Using the slicer states stored configuration file, the rival program uses a python-based algorithm to filter the raw data identically to PowerBI's operations. The filtered underlying data from PowerBI and the rival algorithm are compared in a quality control check to ensure that there are no discrepancies.

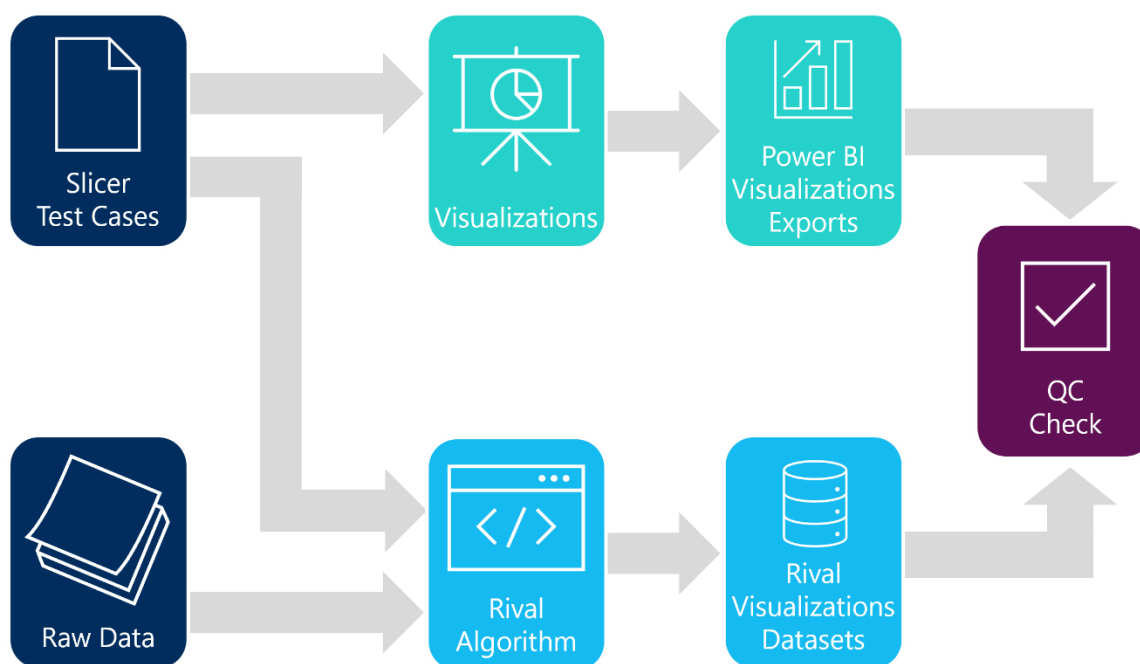


Figure3 – Visualization Validation Process Flow

Limitations

Despite the effectiveness of the proposed validation framework, several challenges remain. One key limitation is the reliance on manual validation steps, particularly in assessing model integrity. While automation can enhance efficiency, certain validation tasks require human oversight to ensure contextual accuracy.

Additionally, the dynamic nature of PowerBI dashboards further complicates validation efforts, as updates and modifications may introduce new variables that require revalidation. These new updates or variables can cause potential issues within rival algorithms or automated processes. As a result, it is necessary to create robust and adaptive programs that can handle potential changes.

Finally, rival programming techniques employ the most recent and cutting-edge software tools; however, there are still some elements of validation that would require software advancements to automate. Unfortunately, there is little that can be done to solve these problems other than waiting for new technologies to be released. Addressing these limitations through continued research and the development of standardized validation protocols will be critical for advancing the reliability and effectiveness of clinical dashboards.

Conclusion

Ensuring data integrity within clinical dashboards is essential for maintaining the accuracy and reliability of healthcare analytics. The three-step validation framework outlined in this paper provides a systematic approach to verifying data ingestion, model integrity, and visualization accuracy. By incorporating automation and rival programming techniques, this methodology enhances quality assurance while reducing manual effort. This approach has the potential to change the way validation is viewed within the pharmaceutical industry. A manually intensive process can be replaced with automated systems that increase confidence in clinical dashboards.

This streamlined form of validation has numerous benefits for the pharmaceutical industry. Decreased time spent on validation and increased assurance of proper data representation can lead to quicker decision making by healthcare professionals and quicker care for patients who could be in desperate need.

Furthermore, the ethical implications of dashboard validation cannot be overlooked. Inaccurate data representation can lead to misinformed clinical decisions that can cause harm to both patients and businesses. Establishing rigorous validation protocols not only safeguards data integrity but also fosters trust in the insights derived from PowerBI dashboards. As clinical dashboards become increasingly relied upon, the need for validation will continue to increase. With the developments of new software, the efficiency of validation tools will continue to increase, and the limitations will decrease.

Moving forward, the integration of industry-wide best practices and continuous refinement of validation techniques will further strengthen the role of dashboards in clinical decision-making, ensuring that they remain valuable tools for data-driven healthcare advancements.

References

1. [Requirements for a quality dashboard: Lessons from National Clinical Audits - PMC](#)

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

River Meckstroth
MMS Holdings
6880 Commerce Blvd
Canton, MI, 48187
Email: rmeckstroth@mmsholdings.com
Website: www.mmsholdings.com

Kris Wenzel
MMS Holdings
6880 Commerce Blvd
Canton, MI, 48187
Email: kwenzel@mmsholdings.com
Website: www.mmsholdings.com

Brand and product names are trademarks of their respective companies.

