

Where Does Your SCE Fit in Your Data Strategy?

Philip Loucks, Instem, North Carolina, United States

ABSTRACT:

A data strategy is akin to a blueprint for managing, interpreting, and utilising data within an organisation. It encompasses plans for data collection, storage, analytics, and alignment with business objectives. Imagine it as a roadmap guiding how data - regardless of its source - should be organised, stored, and accessed. Your statistical computing environment (SCE) is a key step in your data value chain and thus it should feature highly in your organisation's clinical data strategy.

This paper describes the value of a data strategy, shows some example strategies (featuring SCEs), and specifies the most important facets of your SCE's contribution and role in your data strategy.

If you don't already have a data strategy, come and learn how to make your first steps. If you already have a data strategy, come and consider if your SCE is playing its optimal role in the strategy.

INTRODUCTION: WHAT IS DATA STRATEGY?

A data strategy is a comprehensive plan that outlines how an organisation will collect, store, manage, and use its data to achieve its business goals. It's essentially a roadmap for turning raw data into actionable insights.

Key components of a data strategy typically include:

- **Data Governance:** Establishing rules, policies, and standards for data management to ensure its quality, security, and integrity.
- **Data Architecture:** Designing the infrastructure (hardware, software, and processes) for storing, processing, and analysing data.
- **Data Quality:** Implementing measures to ensure data accuracy, completeness, and consistency.
- **Data Security:** Protecting data from unauthorised access, breaches, and misuse.
- **Data Analytics:** Utilising tools and techniques to extract meaningful insights from data.
- **Data Visualisation:** Presenting data in a clear, understandable, and visually appealing format.
- **Data Integration:** Combining data from various sources into a unified view.

Why is a data strategy important?

- **Competitive Advantage:** Data-driven organisations can make more informed decisions, identify new opportunities, and gain a competitive edge.
- **Improved Efficiency:** Data can help streamline operations, reduce costs, and improve customer satisfaction.
- **Risk Mitigation:** A well-defined data strategy can help organisations identify and address potential risks associated with data management.
- **Regulatory Compliance:** Many industries have data privacy and security regulations that must be adhered to.

In essence, a data strategy is about harnessing the power of data to drive business success. A well-defined clinical data strategy helps ensure that clinical trials are conducted efficiently, results are reliable, and valuable insights are gained to advance medical research and patient care.

DATA COLLECTION (SUPPLY)

Data collection is a crucial part of a data strategy, especially in clinical trial data analysis. Here are some key elements involved:

- **Source Identification:** Identifying all potential data sources, such as patient-reported outcomes, laboratory results, and data from wearable devices. This ensures comprehensive data capture.
- **Data Standards:** Establishing standards for data formats, terminologies, and coding systems to ensure consistency and interoperability. This might involve using standardised protocols like CDISC (Clinical Data Interchange Standards Consortium).
- **Data Collection Methods:** Choosing appropriate methods for data collection, such as electronic data capture (EDC) systems, mobile apps, or paper-based forms. The method should align with the trial's needs and regulatory requirements.
- **Data Quality Assurance:** Implementing procedures to ensure data accuracy, completeness, and reliability. This includes training for data collectors, real-time data validation checks, and regular audits.

- **Patient Consent and Privacy:** Ensuring that data collection complies with ethical standards and regulations, such as GDPR or HIPAA. This involves obtaining informed consent from participants and safeguarding their privacy.
- **Data Integration:** Developing processes to integrate data from various sources into a unified dataset. This often requires data transformation and harmonisation to handle different formats and structures.
- **Real-time Data Monitoring:** Setting up systems for real-time monitoring of data collection to identify and address issues promptly. This helps maintain data integrity and allows for timely interventions if needed.
- **Technology and Tools:** Utilising advanced technologies and tools for data collection, such as cloud-based platforms, AI-driven data capture, and blockchain for secure data transactions.

By focusing on these elements, a data strategy can ensure that the data collected during clinical trials is robust, reliable, and ready for analysis, ultimately leading to more accurate and meaningful results

STORAGE

Data storage is a critical component of a clinical trial data strategy, ensuring that all collected data is securely maintained and easily accessible for analysis and reporting. Here are some key aspects:

- **Security and Compliance:** Data storage solutions must comply with regulatory standards such as GDPR, HIPAA, and GCP (Good Clinical Practice). This involves implementing robust security measures to protect sensitive patient information from unauthorised access and breaches.
- **Data Integrity:** Ensuring the integrity of stored data is paramount. This includes using secure, validated systems that prevent data corruption and maintain accurate records over time. Regular backups and audit trails are essential to track any changes or access to the data.
- **Scalability:** Clinical trials can generate vast amounts of data, so storage solutions need to be scalable. Cloud-based storage options are often used because they can easily expand to accommodate growing data volumes and provide flexible access to data from multiple locations.
- **Accessibility:** While security is crucial, authorised personnel must have easy access to the data when needed. This involves setting up user permissions and access controls to ensure that only qualified individuals can view or modify the data.
- **Data Formats and Standards:** Storing data in standardised formats (e.g. CDISC standards) ensures consistency and facilitates data sharing and integration with other systems. This is important for data analysis and reporting.
- **Disaster Recovery:** Implementing disaster recovery plans is essential to protect data against loss due to unforeseen events like natural disasters or system failures. This includes regular data backups and having redundant storage systems in place.
- **Metadata Management:** Properly managing metadata (data about data) helps in understanding the context, provenance, and usage of the stored data. This is crucial for data traceability and quality control.

By focusing on these aspects, a clinical trial data strategy can ensure that data is stored securely, remains accessible to authorised users, and is maintained in a way that supports accurate and reliable analysis.

ANALYTICS (CONSUMPTION):

The analytics element of a clinical trial data strategy is crucial for ensuring the success and integrity of the trial. Here are some key aspects:

- **Statistical Analysis:** Using statistical techniques to evaluate the significance of the findings. This can include identifying patterns and trends, comparing results from different trials, and assessing the efficacy and safety of the treatment or device.
- **Predictive Analytics:** Leveraging historical data to predict future outcomes. This helps in optimising trial operations, such as patient recruitment and retention, and forecasting potential challenges.
- **Real-Time Monitoring:** Implementing systems to monitor data in real-time, allowing for quick adjustments and ensuring the trial stays on track.
- **Data Integration:** Combining data from multiple sources to provide a comprehensive view of the trial. This can include integrating clinical data with genomic data, electronic health records, and other relevant datasets.
- **Regulatory Compliance:** Ensuring that all data analytics processes comply with regulatory requirements to maintain the integrity and credibility of the trial.

The choice of a good statistical computing environment (SCE) is crucial in being able to meet these needs.

BUSINESS ALIGNMENT

Aligning a clinical trial data strategy with business objectives is essential for maximising the value and impact of the trial. Here are some key elements to consider:

- **Stakeholder Involvement:** Engage key stakeholders from the beginning to ensure their needs and perspectives are integrated into the data strategy. This helps in aligning the strategy with the overall business goals.
- **Clear Objectives:** Define clear, measurable objectives that align with the business goals. This includes identifying what the trial aims to achieve and how it supports the broader business strategy.
- **Risk Management:** Given the highly regulated nature of clinical trials, it's crucial to focus on minimising risks and ensuring compliance with regulatory requirements. This alignment helps in maintaining the integrity and credibility of the trial.
- **Data Integration:** Ensure that data from the clinical trial is integrated with other business data sources. This provides a comprehensive view that can inform business decisions and strategies.
- **Performance Metrics:** Establish performance metrics to track the success of the data strategy. These metrics should be aligned with business objectives to measure the impact of the trial on the overall business.
- **Continuous Improvement:** Regularly review and update the data strategy to ensure it remains aligned with evolving business goals and regulatory requirements.

DATA FROM AN SCE CONTEXT

INTRO, WHAT'S AN SCE:

A Statistical Computing Environment (SCE) like Sycamore, Aspire™, and SAS® Life Science Analytics Framework (LSAF) provides a robust platform for managing and analysing data, particularly in clinical research. These environments are essential for ensuring the accuracy, efficiency, and compliance of statistical analyses in clinical trials and other research settings.

A Statistical Computing Environment (SCE) is a centralised, integrated platform designed to streamline and ensure the rigor of statistical analysis processes. Here are some key features and benefits:

- **Data Management:** SCEs provide tools for managing, processing, and analysing large datasets efficiently. This includes version control, data security, and integrity.
- **Automation:** They automate various aspects of the statistical analysis workflow, reducing manual effort and minimising errors.
- **Compliance:** SCEs help ensure that all data analysis processes comply with regulatory requirements, which is crucial in fields like clinical trials.
- **Collaboration:** These environments facilitate collaboration among researchers by providing a shared platform for data and analysis tools.
- **Documentation and Reporting:** SCEs offer capabilities for documenting and reporting the analysis process, which is essential for transparency and reproducibility.

DATA SOURCES:

In a clinical trial, a Statistical Computing Environment (SCE) integrates data from various sources to ensure comprehensive and accurate analysis. Here are some common data sources used in clinical trials:

- **Electronic Health Records (EHRs):** These provide detailed patient information, including medical history, diagnoses, treatments, and outcomes.
- **Clinical Data Management Systems (CDMS):** These systems collect and manage data from clinical trial sites, ensuring data integrity and compliance with regulatory standards.
- **Electronic Data Capture (EDC) Systems:** EDC systems are used to collect data directly from clinical trial participants, often through electronic forms or questionnaires.
- **Laboratory Information Management Systems (LIMS):** These systems manage data from laboratory tests and analyses, ensuring accurate and timely reporting of lab results.
- **Wearable Devices and Sensors:** Increasingly, clinical trials use data from wearable devices and sensors to monitor participants' health metrics in real-time.
- **Patient-Reported Outcomes (PROs):** Data collected directly from patients about their health condition and treatment experiences.
- **Imaging Systems:** Data from medical imaging technologies such as MRI, CT scans, and X-rays, which provide detailed insights into the physiological effects of treatments.
- **Genomic Data:** Information from genetic testing and sequencing, which can be crucial for personalised medicine approaches.

These diverse data sources are integrated within an SCE to provide a holistic view of the clinical trial, enabling robust statistical analysis and ensuring regulatory compliance. Whilst the upstream data management environment (DME) is principally responsible for many elements of data collation and quality, the SCE retains some of these responsibilities too.

DATA CONSUMERS:

Data consumers of a Statistical Computing Environment (SCE) in clinical trials are diverse and play crucial roles in the trial's success. Consumers may make use of presented analytical results or the underlying results data. Here are some key data consumers:

- **Clinical Researchers:** They use the SCE to analyse trial data, identify trends, and draw conclusions about the efficacy and safety of the treatment or intervention.
- **Regulatory Authorities:** Agencies like the FDA and EMA review the data to ensure compliance with regulatory standards and to make decisions about the approval of new treatments.
- **Biostatisticians:** These professionals perform detailed statistical analyses to interpret the data and ensure the validity of the trial results.
- **Data Managers:** They oversee the collection, cleaning, and management of data within their data management environment (DME) and then release the data to the SCE, ensuring data integrity and quality.
- **Medical Monitors:** These professionals review the data to monitor patient safety and manage any adverse events that occur during the trial.
- **Health Economists:** They analyse the data to assess the cost-effectiveness of the treatment, which is crucial for pricing and reimbursement decisions.
- **Academic Institutions:** Researchers and students use the data for educational purposes and to conduct further research.

These data consumers rely on the SCE to provide accurate, timely, and comprehensive data to support their various roles in the clinical trial process.

ANALYTICS:

Analytics in a Statistical Computing Environment (SCE) for clinical trials are essential for ensuring the integrity, accuracy, and efficiency of the trial data. Here are some key types of analytics used:

- **Descriptive Analytics:** This involves summarising and describing the main features of the collected data. It includes generating summary statistics, such as means, medians, and standard deviations, and creating visualisations like histograms and scatter plots to understand data distributions and relationships.
- **Inferential Analytics:** These techniques are used to make inferences about the population based on sample data. This includes hypothesis testing, confidence intervals, and regression analysis to determine the relationships between variables and to test the efficacy of treatments.
- **Predictive Analytics:** Leveraging historical data to predict future outcomes. This can help in identifying potential risks, optimising patient recruitment, and forecasting trial timelines.
- **Prescriptive Analytics:** This type of analytics provides recommendations for decision-making. It uses simulation and optimisation techniques to suggest the best course of action based on the data.
- **Real-Time Analytics:** Implementing systems to monitor data as it is collected, allowing for immediate insights and adjustments. This is crucial for ensuring patient safety and maintaining the trial's integrity.
- **Data Integration and Harmonisation:** Combining data from various sources (e.g., EHRs, EDC systems, wearable devices) to provide a comprehensive view of the trial. This ensures that all relevant data is considered in the analysis.
- **Compliance and Audit Trails:** Ensuring that all data processing and analysis steps are documented and traceable. This is vital for regulatory compliance and for ensuring the reproducibility of results.

In addition to forming the backbone of the Clinical Study Report (CSR) sent to regulators, these analytics processes help in making informed decisions, optimising trial operations, and ensuring the reliability and validity of the trial outcomes.

WHAT DOES A DATA STRATEGY FEATURING AN SCE LOOK LIKE?**OBJECTIVES:**

The objectives of a data strategy for a clinical trial Statistical Computing Environment (SCE) are multifaceted and aim to ensure the trial's success, integrity, and compliance. Here are some key objectives:

- **Data Quality and Integrity:** Ensure that all data stored is accurate, complete, and reliable. This involves rigorous management processes.
- **Regulatory Compliance:** Maintain compliance with all relevant regulatory requirements and guidelines. This includes ensuring that data handling and analysis processes meet standards set by agencies like the FDA and EMA.
- **Efficiency and Automation:** Streamline data processing and analysis workflows through automation. This reduces manual effort, minimises errors, and accelerates the overall trial timeline.
- **Integration and Interoperability:** Facilitate the integration of data from various sources (e.g., EHRs, EDC systems, wearable devices) to provide a comprehensive view of the trial. This ensures that all relevant data is considered in the analysis.

- **Real-Time Monitoring and Analytics:** Implement real-time data monitoring and analytics to enable timely insights and adjustments. This is crucial for maintaining patient safety and ensuring the trial stays on track.
- **Transparency and Traceability:** Ensure that all data processing and analysis steps are transparent and traceable. This is vital for regulatory audits and for ensuring the reproducibility of results.
- **Stakeholder Engagement:** Engage all relevant stakeholders, including clinical researchers, biostatisticians, regulatory authorities, and data managers, to ensure their needs and perspectives are integrated into the data strategy.
- **Scalability and Flexibility:** Design the data strategy to be scalable and flexible, allowing it to adapt to the evolving needs of the trial and future studies.

These objectives help in creating a robust data strategy that supports the successful execution and analysis of clinical trials.

OPTIONS

When developing a data strategy for a clinical trial Statistical Computing Environment (SCE), several options and approaches can be considered to ensure efficiency, compliance, and robust data management. Here are some key options:

- **Integrated Platforms:** Utilise integrated e-clinical platforms that combine various technologies for managing clinical trial planning and execution. These platforms provide access to metadata, clinical data, and operational data, streamlining workflows and enhancing data accessibility.
- **Automation and Workflow Management:** Implement automation tools to streamline data processing and analysis workflows. This includes scheduling programs to execute automatically upon data updates, which increases efficiency and reduces manual errors.
- **Data Governance and Compliance:** Establish strong data governance frameworks to ensure compliance with regulatory requirements. This includes controlled access, version control, audit trails, and traceability of data and analysis processes.
- **Real-Time Data Monitoring:** Incorporate real-time data monitoring systems to enable immediate insights and adjustments. This is crucial for maintaining patient safety and ensuring the trial stays on track.
- **Scalability and Flexibility:** Design the SCE to be scalable and flexible, allowing it to adapt to the evolving needs of the trial and future studies. This includes supporting various data sources and types, such as EHRs, wearable devices, and genomic data.
- **Collaboration Tools:** Use collaboration tools within the SCE to facilitate communication and data sharing among researchers, biostatisticians, and other stakeholders. This enhances teamwork and ensures that all relevant expertise is leveraged.
- **Advanced Analytics:** Leverage advanced analytics, including predictive and prescriptive analytics, to optimise trial operations and make informed decisions. This can help in identifying potential risks, optimising patient recruitment, and forecasting trial timelines.
- **Data Integration and Harmonisation:** Ensure seamless integration and harmonisation of data from various sources to provide a comprehensive view of the trial. This includes integrating clinical data with genomic data, electronic health records, and other relevant datasets.

These options help in creating a robust and effective data strategy for clinical trial SCEs, ensuring that the trial is conducted efficiently, compliantly, and with high data integrity

ADOPTING A DATA STRATEGY:

CHANGE MANAGEMENT:

Adopting a data strategy for a clinical trial Statistical Computing Environment (SCE) involves significant change management to ensure a smooth transition and successful implementation. Here are some key aspects of change management in this context:

- **Stakeholder Engagement:** Involve all relevant stakeholders early in the process. This includes clinical researchers, data managers, IT staff, regulatory experts, and senior management. Their input and buy-in are crucial for the strategy's success.
- **Clear Communication:** Develop a clear communication plan to keep all stakeholders informed about the changes, the reasons behind them, and the benefits. Regular updates and open channels for feedback are essential.
- **Training and Support:** Provide comprehensive training and ongoing support to ensure that all users are comfortable with the new systems and processes. This can include workshops, online tutorials, and help desks.
- **Pilot Testing:** Implement the data strategy in phases, starting with a pilot test. This allows you to identify and address any issues before a full-scale rollout.
- **Risk Management:** Identify potential risks associated with the change and develop mitigation strategies. This includes technical risks, such as data migration issues, and human factors, such as resistance to change.

- **Monitoring and Evaluation:** Continuously monitor the implementation process and evaluate its effectiveness. Use metrics and feedback to make necessary adjustments and improvements.
- **Cultural Change:** Foster a culture that embraces change and innovation. Encourage a mindset that views the new data strategy as an opportunity for improvement rather than a disruption.
- **Documentation and Compliance:** Ensure that all changes are well-documented and comply with regulatory requirements. This is crucial for maintaining data integrity and meeting audit standards.

By focusing on these aspects, you can effectively manage the transition to a new data strategy for your clinical trial SCE, ensuring that it enhances efficiency, compliance, and data quality.

GOVERNANCE:

Adopting a data strategy for a clinical trial involves establishing a robust governance framework to ensure data integrity, security, and compliance. Here are some key aspects to consider:

- **Data Governance Structure:** Establish a clear governance structure that includes a data management group reporting to the data strategy owner. This structure should have the authority to enforce the data strategy.
- **Alignment of People, Processes, and Technology:** Data governance aligns people, processes, and information technology to optimise the use and value of data. This helps in collecting, integrating, and analysing data strategically to advance drug development programs.
- **Compliance and Integrity:** Ensure compliance with regulatory guidelines such as ICH E6 (R3), which emphasises data integrity, traceability, and security. This includes safeguarding participant confidentiality and maintaining accurate audit trails.
- **Data Management Practices:** Implement best practices for data management, including data blinding to minimise bias, and regular audits to assess data integrity.
- **Training and Competency Development:** Develop data science competencies within the team to handle the increasing volumes and complexity of clinical trial data.

By focusing on these areas, you can create a data governance framework that supports efficient and compliant clinical trial operations.

FUTURE CHALLENGES AND OPPORTUNITIES

AI/ML/GEN AI

Integrating AI, ML, and Generative AI (GenAI) into clinical data strategies presents both significant opportunities and challenges. Here's a breakdown:

Opportunities:

- **Enhanced Data Analysis:** AI and ML can process vast amounts of clinical data quickly, identifying patterns and insights that might be missed by human analysts. This can lead to more accurate predictions and better decision-making in clinical trials.
- **Personalised Medicine:** GenAI can help in developing personalised treatment plans by analysing patient data to predict individual responses to treatments. This can improve patient outcomes and reduce adverse effects.
- **Efficiency in Clinical Trials:** AI and ML can streamline various aspects of clinical trials, from patient recruitment to data collection and analysis. This can reduce the time and cost associated with bringing new drugs to market.
- **Improved Patient Monitoring:** AI-powered tools can continuously monitor patient data, providing real-time insights and early warnings about potential issues. This can enhance patient safety and trial integrity.

Challenges:

- **Data Privacy and Security:** Handling sensitive patient data requires stringent security measures to prevent breaches and ensure compliance with regulations like GDPR and HIPAA.
- **Bias and Fairness:** AI and ML models can inherit biases present in the training data, leading to unfair or inaccurate outcomes. Ensuring fairness and transparency in AI models is crucial.
- **Regulatory Compliance:** The use of AI in clinical trials must comply with regulatory standards, which can be complex and vary by region. Navigating these regulations requires careful planning and expertise.
- **Integration with Existing Systems:** Incorporating AI and ML into existing clinical trial infrastructure can be challenging. It requires significant investment in technology and training for staff.
- **Data Quality and Management:** Ensuring high-quality, standardised data is essential for effective AI and ML applications. Poor data quality can lead to inaccurate models and unreliable results.

By addressing these challenges and leveraging the opportunities, clinical trials can become more efficient, accurate, and patient-centric.

ACTIONABLE INSIGHTS / CONCLUSION

Integrating a Statistical Computing Environment (SCE) into a clinical trial data strategy can significantly enhance data management and analysis. Here are some actionable insights to help you fit an SCE into your clinical trial data strategy:

1. Define Clear Objectives

- **Identify Goals:** Determine what you aim to achieve with the SCE, such as improving data analysis efficiency, ensuring compliance, or enhancing data security.
- **Set Metrics:** Establish key performance indicators (KPIs) to measure the success of the SCE integration.

2. Ensure Regulatory Compliance

- **Adhere to Guidelines:** Ensure the SCE complies with regulatory standards like FDA 21 CFR Part 11, GDPR, and HIPAA.
- **Audit Trails:** Implement robust audit trails to track data access and modifications.

3. Data Integration and Management

- **Centralised Data Repository:** Use the SCE to create a centralised repository for all clinical trial data, ensuring consistency and accessibility.
- **Data Standardisation:** Standardise data formats and structures to facilitate seamless integration and analysis.

4. Enhance Data Security

- **Encryption:** Implement encryption for data at rest and in transit to protect sensitive information.
- **Access Controls:** Establish strict access controls to ensure only authorised personnel can access the data.

5. Leverage Advanced Analytics

- **AI and ML Integration:** Utilise AI and ML tools within the SCE to analyse data more effectively and uncover deeper insights.
- **Real-Time Monitoring:** Implement real-time data monitoring to quickly identify and address any issues.

6. Training and Support

- **Staff Training:** Provide comprehensive training for staff on how to use the SCE effectively.
- **Ongoing Support:** Ensure there is ongoing technical support to address any issues that arise.

7. Continuous Improvement

- **Feedback Loop:** Establish a feedback loop to continuously gather input from users and stakeholders to improve the SCE.
- **Regular Updates:** Keep the SCE updated with the latest features and security patches.

By following these actionable insights, you can effectively integrate an SCE into your clinical trial data strategy, enhancing data management, compliance, and analysis capabilities.

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

Name: Philip Loucks

Company: Instem

Work Phone: +1 (267) 664-3713

Email: phil.loucks@instem.com

Website: www.instem.com

Brand and product names are trademarks of their respective companies.