DIGITAL HEALTH TRIALS

Recommendations for Managing Data

Digital health technologies have fundamentally changed when, where, and how data can be collected. Yet the central principle of data collection remains the same: ensuring the authenticity, integrity, and confidentiality of data over time.

CTTI's recommendations and resources below can help you understand, plan for, andaddress the new challenges associated with digital health technologies, including data access, ownership, and storage; data sharing considerations; and communicating trial data with participants.

This document provides a set of high level recommendations. A detailed considerations document is available under the Resource section of this document and at https://ctti-clinicaltrials.org/our-work/digital-health-trials/managing-data/

RECOMMENDATIONS

General

1. Collect the minimum data set necessary to address the study endpoints

- Quality by Design principles should drive decisions about the quantity of data to be collected.
- When the study endpoints are well understood, CTTI recommends against speculative "data fishing."
- Collect applicable metadata, where appropriate, to allow statisticians and data scientists to identify data that may be irregular, inconsistent, or confounded during data cleaning prior to analysis.

2. Identify acceptable ranges and mitigate variability in endpoint values collected via digital health technologies

- Eliminate sources of variability in data quality, for example, by reducing variability in measurements.
- The most appropriate epoch length and optimal sampling frequency for a given outcome should be determined during endpoint development according to the context of use within the trial.

Data Access and Interpretation

3. Optimize data accessibility while preventing data access from unauthorized users

- The security principles of "*need to know*" and "*least privilege*" should be applied when determining access rights and privileges.
- Rights and privileges should be regularly reviewed and updated to prevent risks to the integrity, privacy, and confidentiality of data.

4. Ensure that the data access plan meets your needs prior to contracting an electronic service vendor

- Sponsors should consider two primary factors before entering into an outsourcing agreement with a digital technology manufacturer:
  - Whether data generated by the digital health technology may be accessed and used by the manufacturer, and
  - What data will be provided by the manufacturer to the sponsor.

5. Address data attribution proactively with patient input

- A multi-pronged approach should be used to promote correct ***data attribution***. Strategies should focus on the following considerations:
  - Digital health technology selection
  - Protocol design
  - Technical approaches
- Sponsors should engage participants in decisions regarding the inclusion of technical approaches to ensure correct data attribution.

6. Ensure that site investigators have access to data generated by their participants

- CTTI recommends that read-only application programming interfaces (API) take the form of a dashboard with data summarized at the appropriate levels of detail for the investigators' needs.

7. Identify ways to return value to participants throughout the trial, including the return of outcomes data collected by digital health technologies

- During the trial planning process, carefully review the outcomes and other health-related information that will be generated during the trial and develop a plan for returning data to participants, including when and how it will be returned.
- Planning should incorporate patient and site perspectives and should provide participants with appropriate access to their individual data as well as aggregate trial results.
- Provide participants with real-time access to their individual outcomes or other data only if it can be done in a way that maintains study integrity and participant safety.
- When returning data to participants, do so in a way that is both personalized and understandable.
- Clearly communicate plans for sharing data/results with potential participants during the informed consent process.

8. Let data sharing decisions be driven by safety and trial integrity.

- Investigators should consider sharing summarized data with each individual at the completion of their enrollment as well as offering study participants the opportunity to learn about the overall results of the trial when it is completed.
- The ways in which data generated by digital health technologies will be used within the trial as well as how this data may be accessed and used by the technology manufacturer and any additional third parties should be clearly stipulated in the outsourcing agreements, and a clear accounting of which parties will have access to each level of data should be included in the informed consent and HIPAA research authorization form.
- Sponsors should engage potential participants in discussions regarding access to and use of data by external entities to reach a decision that ultimately meets the patients' levels of comfort and expectations of privacy.
- To promote collaboration through data sharing, CTTI recommends that all stakeholders reference the Institute of Medicine's 2015 report, "Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk."[1]

9. Establish industry-wide standards to drive the successful scaling and accelerated acceptance of clinical trials using digital technologies for data capture.

- CTTI recommends establishing industry-wide standards related to the following digital health trial considerations:
    1. Terminology—for example, the definition of "*raw data*" versus "analysis-ready data;"
    2. The collection and reporting of data captured by digital technologies, including metadata;
    3. Transparency of information related to digital technology specifications, calibration, and verification bench-tests; and
    4. Transparency requirements for the development of algorithms used to convert the data into physiologically and medically useful endpoints.

Security & Confidentiality

10. Apply an end-to-end, risk-based approach to data security

- An end-to-end, risk-based approach to data security should be applied to protect participants' privacy and the confidentiality and integrity of their data.
- When taking a risk-based approach to data security, sponsors should expect outsourced electronic service vendors to conduct comprehensive security assessments prior to developing their risk-based security solutions.
- Security assessments, including those conducted by sponsors implementing their own security solutions, should include these six domains: dependencies on outside providers, systems, procedures, people, policies, and applicable regulations.
- Prior to signing an outsourcing agreement with an electronic service vendor, as with any outsourced information system, the sponsor and relevant audit or compliance teams should review and evaluate the vendor's security systems.

11. Ensure the authenticity, integrity, and confidentiality of data over its entire lifecycle

Data authenticity means that the data are what the originator claims the data to be.

- For outcomes data captured by a digital health technology to be authentic, the data should display all of the characteristics outlined below:
    - Precise and accurate representation of what the data claim to be measuring—for example, heart rate (bpm) or activity (steps/day)
    - Correct attribution of data to the intended participant
    - Inclusion of metadata indicating the source of the data and a UTC time stamp
    - Demonstration that the data have not been corrupted following creation
- CTTI recommends addressing the integrity of data generated using digital health technologies during both trial planning and conduct. This should include clarifying, at each step during the data lifecycle, the following considerations:
    1. Who is responsible for the data?
    2. Who is accountable for data integrity?
- Sponsors should develop strategies for maintaining the confidentiality of data collected by digital health technologies in concert with developing strategies for protecting participants' privacy.
- To maintain confidentiality, study participants should be informed of and consent to the ways in which their identifiable, private information will be handled.
- Care should be taken to catalog all HIPAA identifiers that a digital health technology collects, even for identifiers that are not the focus of the protocol.

Monitoring for Safety & Quality

12. Include appropriate strategies for monitoring and optimizing data quality

- Data should be collected by digital health technologies in such a way as to optimize the quality of the data.
- Sponsors should design trials to ensure that the use of digital health technologies for data collection does not give rise to new data quality issues. Specifically, that the clinical trial design ensures data captured from digital health technologies are accurate, complete, and correctly attributed.

13. Communication and transparency with participants regarding safety monitoring is critical

- Study participants should be well informed regarding the level of safety monitoring, if any, that will occur when they use/wear their digital health technology.
- Information about the level of safety monitoring must be described in the informed consent.
- If a digital health technology is intended to be used for detecting a safety signal or adverse event, this specific measure should be valid and well understood in the intended context of use.
- When a digital health technology is relied upon to accurately detect a pre-specified safety signal, measures recorded outside of acceptable limits should be directly communicated to the investigators and sponsors via automated processes and algorithms.
- Any actions to be taken following the detection of a safety signal or adverse event by a digital health technology should be pre-specified in the protocol and clearly communicated to study participants.
- Study participants should be included in decision-making on how to handle responses to safety signals and adverse events detected remotely by digital health technologies.

14. Plan appropriately for the statistical analysis of data captured using digital technologies

- When the research question caters to the use of digital health technologies for the collection of large data sets, sponsors should ensure access to both suitable data platforms to handle this volume of data and the necessary expertise to manipulate it at different levels of granularity.
- Regardless of the anticipated size of the data set, sponsors should consider conducting small-scale feasibility studies prior to finalizing their protocol design to ensure familiarity with the nature of the data outputs from the digital health technology(ies) and the correct analytical approach.
- Statistical analysis plans should be fit for purpose and developed prior to trial initiation.
- Biostatisticians should be involved in planning analyses during protocol

development.

15. Monitor data quality centrally through automated processes

- When digital health technologies are used for data capture, existing guidance describing strategies for monitoring data quality still applies.
- When digital health technologies are included in the protocol, the quality—specifically the completeness, consistency, and correctness—of the data captured should be monitored centrally.
- To maintain data security and privacy and to promote efficiency, programming and algorithms should be the preferred techniques for verifying data quality, with programmed alerts sent when potential issues are identified.
- Sponsors should consider a centralized approach to monitoring data collected by digital health technologies to track other aspects of trial conduct and progress.
- As monitoring should not only detect deficiencies in trial conduct but also strive to correct them, data monitoring plans should clearly state who is expected to take appropriate action in response to potential issues identified.

16. Minimize missing data

- When using digital health technologies for data capture, a multi-pronged approach to preventing missing data is optimal, with efforts focused on the following strategies:
    1. Optimizing trial design,
    2. Ensuring that established technical approaches eliminate any technology- or transmission-related causes of missing data, and
    3. Pilot testing to identify any unanticipated causes of missing data.
- If a sponsor selects a digital health technology that relies on a companion app for data transfer, the sponsor should ensure that this app is capable of pairing with the wearable technology or remote sensor with minimal risk of data loss.
- For those technologies that connect to manufacturers' servers before providing data to the sponsor via a third-party interface, appropriate syncing strategies are critical, as digital health technologies may limit or lose data after periods of non-syncing.

RESOURCES

- Considerations for Advancing the Use of Digital Technologies for Data Capture & Improved Clinical Trials Document
- Digital Health Technologies Data Flow Diagram
- Decision Support Tool: Real-Time Data Sharing with Study Participants
- Table: Approaches for Safety Monitoring & Managing Safety Signals when using Digital Health Technologies for Data Capture
- Table: Promoting & Protecting Data Integrity
- Table: CTTI Recommended Strategies for Optimizing Data Quality
- Case Study: Returning Value to Participants without Compromising StudyIntegrity
- Case Study: Sharing Data to Promote Patient Engagement
- Case Study: Using Remote, Smartphone-Based Data Collection to Share Health Insights
- Case Study: Optimizing Data Quality and Participant Privacy
- Glossary: CTTI's Digital Health Technologies Recommendations

REFERENCES

1. Institute of Medicine (US). Sharing clinical trial data: maximizing benefits, minimizing risk. Washington, DC: The National Academies Press; 2015. Available from: https://www.nap.edu/catalog/18998/sharing-clinical-trial-data-maximizing-benefits-minimizing-risk.