Journal of the Society for Clinical Data Management

Chin Fatt CR, et al. Research Electronic Data Capture Using REDCap and External Data Quality Pipeline Development. *Journal of the Society for Clinical Data Management*. 2025; 5(1): 11, pp. 1–7. DOI: https://doi.org/10.47912/jscdm.431

ORIGINAL RESEARCH

Research Electronic Data Capture Using REDCap and External Data Quality Pipeline Development

Cherise R. Chin Fatt*, Ryan Y. Becker*, Lorraine Burkhalter*, Brennan Dawson*, Thomas J. Carmody*,† and Madhukar H. Trivedi*

Objective: Research Electronic Data CAPture (REDCap) is a powerful web-based data management tool commonly used in academic research centers. Proper use, development, and execution of REDCap are crucial in facilitating high-quality data capture. While REDCap provides integrated tools for assessing data quality and status, these often fail to address the intricate demands of multisite longitudinal studies.
Methods: We present a framework to optimize REDCap project development and introduce a Python-based data quality pipeline.
Results: By focusing on strategic pre-production project design and implementing rapid-response quality assurance during post-production, using an in-house built quality control platform, we substantially improved the quality and accessibility of data. The Blackbox was first released in November 2024 as a platform to identify issues pertaining to data entry (such as errors in numeric data entered as text), data quality (such as incorrect ranges and logic issues), and data integrity (such as missing values). In the first execution of the Blackbox on a clinical trial, 1949 potential data errors were identified. Most of these errors were due to updates to the study's protocol or due to missing branching logic. The study team addressed and resolved the errors and corrected the fields with missing branching logic. Blackbox execution was applied against the corrected dataset. A comparison between the two runs showed that all data issues were resolved.
Conclusion: The REDCap design workflow and quality control platform can empower researchers to enhance both the accuracy and usability of data in complex research projects by leveraging REDCap's capabilities.

Keywords: REDCap; Electronic Data Capture; Quality Control; Python; Blackbox

1. Introduction

Data capture has been a vital component of research studies, from the days of paper and pencil, through manual entry using Microsoft Excel, to sophisticated Web-based applications used to collect data in real-time. REDCap (Research Electronic Data CAPture) is a browser-based electronic data capture (EDC) system developed in 2004 at Vanderbilt University to aid in data capture for clinical and translational studies.[1] Over 5,900 institutional partners worldwide use REDCap. It is cost-effective, includes training resources, is user-friendly, versatile, has good data management, and allows integration across different platforms.[1]

The web portal features an intuitive developer-facing interface, allowing project staff to create electronic case report forms (eCRF) with minimal programming expertise. Current features include support for conditional logic, skipping logic tailored to study subjects, subject-specific surveys, automated alerts, data extraction and reporting, multisite data management, and audit capability, all within a Health Insurance Portability and Accountability Act (HIPAA) and 21 Code of Federal Regulations (CFR) compliant secure system.

Ensuring data cleanliness and quality is a crucial aspect of any research study. Data cleanliness refers to data that is free from errors or inconsistencies which could compromise its reliability or usability. Clean data is well-organized, properly formatted, and devoid of duplicates and irrelevant information, with low rates of missing values. REDCap includes a built-in workflow for data cleaning using the Data Resolution Workflow that

* Center for Depression Research and Clinical Care, Peter O'Donnell Jr. Brain Institute and Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA

† Department of Health Data Science and Biostatistics, Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas, Texas, USA

Corresponding author: Madhukar H. Trivedi, M.D., (madhukar.trivedi@utsouthwestern.edu)

identifies potential data errors. After manually reviewing a potential data error, the study personnel complete a protocol deviation, correct the data error, or justify the data entry within REDCap. Although this process may be suitable for small projects, it becomes burdensome for longitudinal and multisite observational trials involving hundreds to thousands of enrolled participants.

Although the REDCap Data Resolution workflow tool provides adequate functionality for potential data error identification, resolution, and management, as well as robust audit capabilities, it falls short in several critical areas. The tool lacks the capacity to handle complex logic and is not scalable for large datasets. Additionally, the existing data architecture limits customization, particularly in identifying missing information. For instance, the system does not save missed visits where no forms are completed in the dataset, preventing them from being flagged as a data error. This limitation extends to the reporting interface, where reports of incomplete data can be undercounted.

At the Center for Depression Research and Clinical Care (CDRC) at the University of Texas Southwestern Medical Center (UTSW), REDCap is our primary EDC system. Over the past decade, our experience with REDCap has yielded valuable insights into designing workflows for accurate data capture, increasing efficiencies, and streamlining reporting processes. By optimizing project pipelines, refining methodologies, and incorporating technological advancements, we have strengthened data integrity and ensured compliance across a wide range of studies. Thus, the purpose of this manuscript is to share our experience with (1) developing a robust REDCap project pipeline designed for efficient and effective data capture, and (2) establishing a quality control pipeline for data collected using REDCap. These implementations have yielded user-friendly EDC systems, clean data, and streamlined processes for regulatory and compliance reporting.

2. Methods

2.1 REDCap Project Development Overview

Over the past decade, the CDRC has learnt various lessons from using REDCap as our primary EDC system. Some of these lessons include (1) use of consistent eCRFs across projects, (2) thorough project testing, (3) moving projects from development to production for audit purposes, and (4) careful review of changes to projects once data entry has started. These lessons have led us to develop a REDCap project development workflow **(Figure 1)**.

Developing an efficient and effective REDCap project is a collaborative effort between the study team and the data team **(Figure 1)**. Once a new REDCap project is needed or edits are required to an existing project, a request is submitted to the data team for review. The data team begins by reviewing the submitted documentation, which includes the study protocol, study design, timeline, workflow, and any newly developed instruments. If clarification is needed, the data team meets with the study team; otherwise, development begins. Upon completion of the development project, the team takes the following steps: (1) thorough testing by the study team (independent from the data team), (2) ensuring clean and clear formatting of forms, (3) defining user rights and roles with precision, and (4) implementing a comprehensive plan for quality control. Once project development is complete and approved for release by the study and data teams, the project is moved to REDCap production status. This status change provides additional safety through user rights and roles, ensuring only the data team can modify any part of the project. Internally, after a project moves to production, the team creates a full copy. Data management strictly limits access to this copy to ensure continuity. The team limits changes to the first six months of use. This approach ensures proper system utilization while preventing spontaneous changes, a common cause of disorganized data, data loss, and inconsistent data.

Clean and clear formatting of forms in REDCap is essential to the front-end user. We therefore ensure that, for subject-facing questionnaires, forms are easy to read, instructions are clearly labeled, free text fields are validated (e.g., only numbers), and alerts (e.g., emergency numbers or guidance in the event of a medical situation) are clearly identified. This involves data validation for fields (e.g., numeric-only fields), HTML formatting for instructions and alert fields, and hiding fields that are used only by the research team (e.g., date completed and scoring).

Adding REDCap alerts is essential for notifying the study team if a participant may need medical attention (e.g., suicidal or with an out-of-range lab value) rather than manual checking of individual REDCap records. This expedites notification via email with the necessary information and a link to the REDCap record promoting subject safety.

Once a robust REDCap project has been built, the next important aspect is data integrity. Depending on the size and complexity of the project, REDCap's built-in Data Resolution Workflow is sufficient. However, as discussed earlier, this workflow may not be optimal or sufficient for longitudinal and/or multisite projects.

2.2 Blackbox Overview

With the launch of several multisite and longitudinal projects, we realized that the built-in REDCap Data Resolution Workflow could not handle complex logic, was not scalable for large datasets, and was limited in its ability to identify missing information. A simplified version of the Blackbox was built and executed, which showed great potential in error identification and reporting. This simplified version of the Blackbox was the foundation of the CDRC's Blackbox.

The CDRC's Blackbox is a tool that identifies data quality issues **(Figure 2)**. Once corrected, the high-quality data are available for dashboards, publications, and grant applications. The Blackbox is built using Python 3.10, with additional required input documents, such as the study's assessment schedule, visit range (± day), data dictionary, and required measures from the study's protocol, included in format files. The Blackbox automates and standardizes data quality processes across diverse clinical

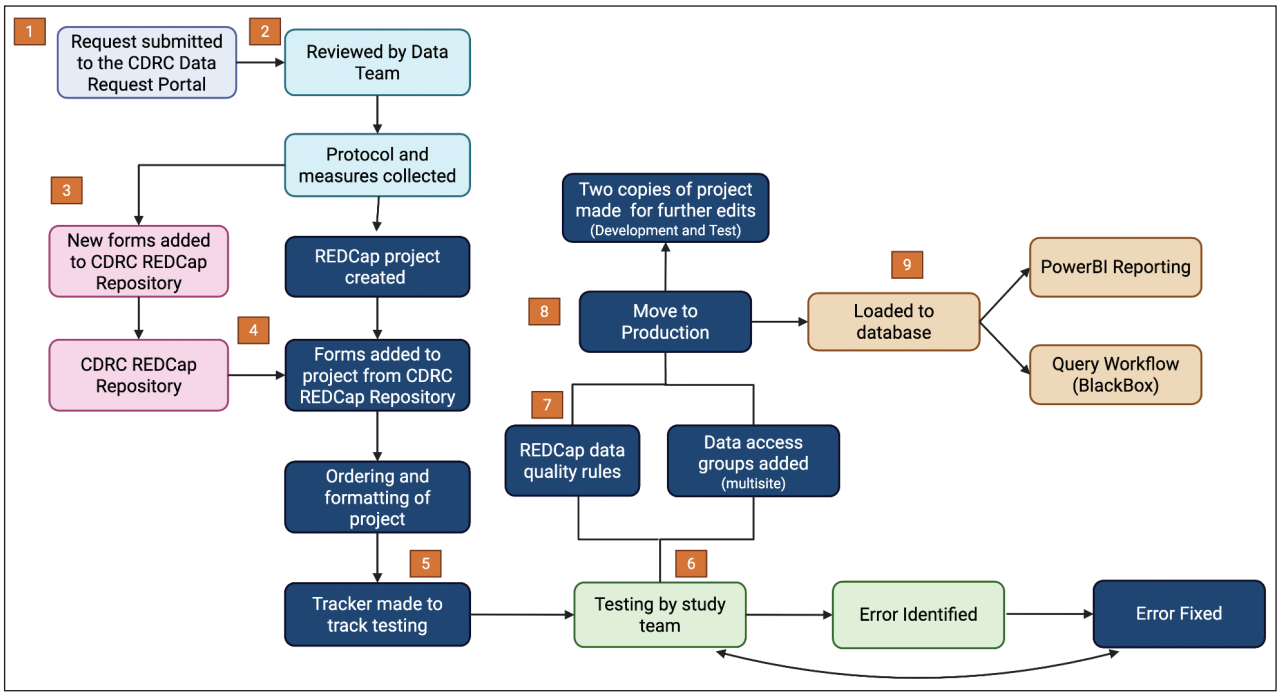

Figure 1: The currently implemented REDCap development pipeline. The different colors within the figure indicate the components of the pipeline, namely (1) a request submitted to the Data Request Portal for the development of a new REDCap project; (2) a request reviewed by the data team, and documents collected from study team; (3) forms added to the internal REDCap repository; if the form already exists in the repository then no action is needed; (4) forms added, ordered and formatted to new REDCap project; (5) testing activity tracker; (6) testing completed by the study team and errors fixed by the developer; (7) internal REDCap data quality rules and data access groups for multisite studies (if applicable) added; (8) project moved to production status and two copies made for future edits; and (9) scripts prepared and loaded to database once one participant has been recorded. This data is then used for reporting and Blackbox query resolution.

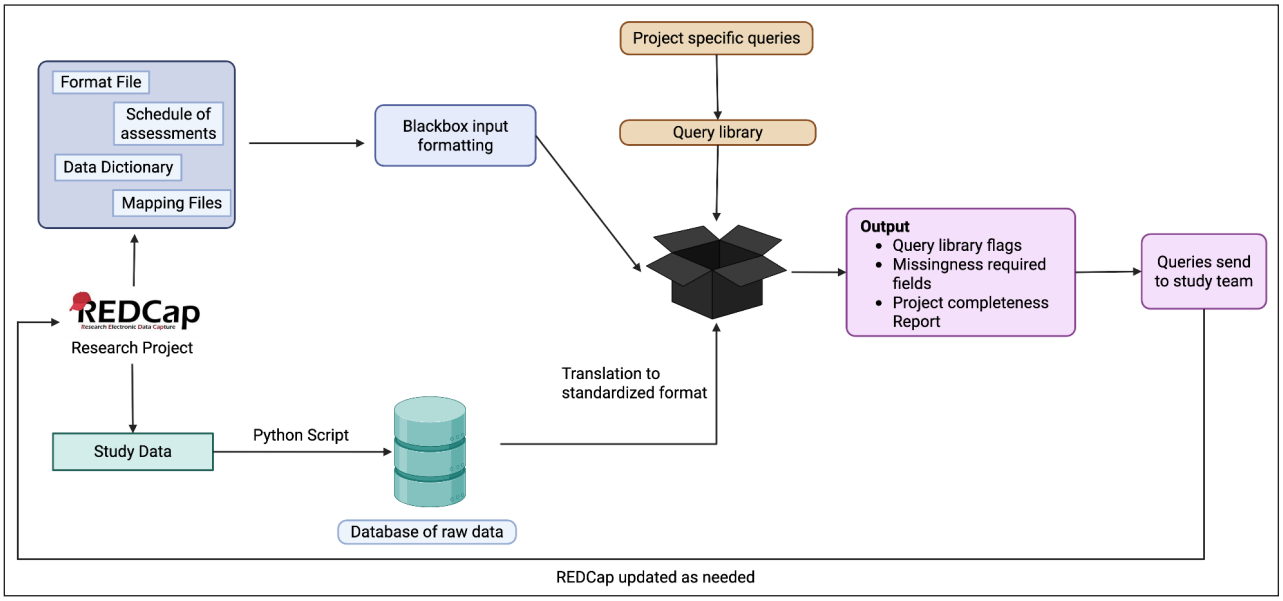

Figure 2: Blackbox Schema. Input documents (blue box) include the format file, schedule of assessments, data dictionary, and mapping files. These documents are formatted for use within the Blackbox. Study data (green) are extracted from REDCap using a Python script and loaded into a SQL database. The translated data is ready for use within the Blackbox. Queries (orange) and the Blackbox input are run against the data and processed within the Blackbox. The output (pink) includes query library flags, missingness required fields, and project completeness report. The study team receives output for resolution. Scheduled Blackbox runs assess new input to identify new issues.

and observational projects, eliminating the need for new project data-quality pipelines. As data collection continues and new projects are launched, identifying data quality issues allows for quick resolution.

A key feature of the Blackbox is its ability to utilize the project's data dictionary from REDCap to determine the fields required across study visits. Using the REDCap data dictionary, the Blackbox scans all expected 'required' fields across study visits. For example, a data error occurs if data is required but no data is entered, but there is no data error if data is not required and not entered. Due to the complexities of branching logic, manual auditing often results in both overlooked fields and incorrect assessments of certain missing data. Blackbox thus reviews the data, applies branching logic within the context, flags the input as necessary, and consistently and accurately identifies issues for evaluation.

An area of particular importance is the ability to report protocol deviations. Items that qualify as a protocol deviation are documented in the study's protocol. One such example is missed visits or forms, which in some studies are considered protocol deviations. Tracking these deviations is essential for regulatory reporting purposes. Accurate reporting of protocol deviations simplifies the data quality pipeline and ensures regulatory compliance. The study team is responsible for reporting these in the protocol deviation instrument in REDCap. Protocol deviations primarily include (1) missed visits, (2) missed forms, (3) missed fields, and (4) incomplete forms, depending on the nature of the study. The protocol deviation form with REDCap provides the Blackbox with all the necessary information to assess study needs. Based on these needs, the Blackbox identifies whether a record of protocol deviations exists or detects newly arising deviations. It further detects and reports incomplete 'Missing Fields' or 'Missed Forms' protocol deviations where documentation does not exist in the protocol deviation form. Once corrected and reported, the Blackbox ceases to flag the violation. This adaptive functionality of the Blackbox, which dynamically examines query resolution statuses through protocol deviations, significantly enhances team efficiency by eliminating the need for labor-intensive updates and manual resolutions.

2.3 Blackbox Input

As previously discussed, the input information for the Blackbox is derived from the project's protocol and from REDCap. The data team manually completes the input files and places them in the designated study folder. Aside from the first-run initialization, the Blackbox accesses these files for instructions and data. This streamlined processing allows the Blackbox to run mostly 'offline'. Further, this compartmentalization avoids unnecessary burdens on the institution's REDCap's server by storing information where required for future runs rather than fetching it live at execution.

The first required input to the Blackbox is the format file, which addresses the problem of the Blackbox striving to work across diverse projects with varying protocols and REDCap designs. The format file codifies information from the project's protocol, such as study flow, consenting processes (in-person or electronic consent), eligibility procedures, what warrants classification as a protocol deviation, and the REDCap project, including variable and event names for critical fields. It also facilitates quality control over repeating forms and events, which can pose problems in data quality pipelines due to how REDCap organizes this data. As long as a REDCap project is designed following specific, robust principles for project design, the format file can successfully encapsulate the complex needs of that project.

Before the Blackbox's first run on a project, the data team directly exports two other Blackbox requirements: (1) the schedule of assessments, and (2) the data dictionary from REDCap. Minimal manual formatting is done on both documents. For the schedule of assessments, a group identifier identifies different schedules based on participants' group membership, such as healthy controls versus patients. The Blackbox manages this by multiple schedules of assessment files. The visit window information allows the Blackbox to identify data collected outside the window.

A lengthy series of regular-expression transformations applied to the data dictionary after acquisition translate REDCap branching logic into Blackbox-compatible syntax, which allows the Blackbox to examine data accurately on a field-by-field level with context from the rest of the record at data collection. In adhering to compartmentalization where possible, the Blackbox does not re-acquire this data from REDCap, unless project changes occur during production.

Additional required inputs to the Blackbox only include information about the language (e.g., English and Spanish translation) and age (e.g., forms specific to pediatric and adult participants). Other types of logic (such as sex specific questions) are handled by branching logic for the specific field or form. Where multiple versions of a given form exist (e.g., forms specific to pediatric and adult populations), multiple mapping files are necessary. To correctly identify violations, the mapping explains the relationship between these forms, understanding that only one of these two forms is expected at a given point in time. An age mapping file informs the Blackbox, which forms links to age-related factors, and also informs how the completion behavior of parent/guardian/legally authorized representative (LAR) forms may change based on the ages of the participants. The language mapping file groups alternate language versions of each form, allowing Blackbox to inspect the project for data in any language. Both the age mapping and language mapping files are optional and are provided as needed.

Independent from the project-specific input files, the Query Library is a repository of forms and associated data rules that provide flexibility and control over data management. As new issues arise during a project's life, rules to identify and flag those problems are seamlessly updated within the Query Library and are executed the next time the Blackbox runs. This approach enables quick additions to the REDCap Data Resolution, addressing limitations such as the inability to detect

certain missed visits, all without needing to edit the Blackbox codebase.

At execution, the Blackbox fetches the most recent version of the project's data from a Structured Query Language (SQL) database and saves it locally to the affiliated folder. In projects where no SQL database exists, the Blackbox initiates a direct application programming interface (API) pull of the most recent data from REDCap. However, this approach may significantly increase runtime and raise loads on the institution's REDCap server.

2.4 Blackbox and Data Pipeline Process

REDCap's built-in Data Resolution Workflow is sufficient for most projects, but becomes unwieldy when scaled. After manually reviewing a query, the study personnel either complete documentation for a protocol deviation or address the issue; they then enter the Data Resolution Workflow to comment on the resolution and notify the data coordinator for review, feedback, and resolution.

The Blackbox's data pipeline is a substantial improvement in both efficiency and accuracy over previous data-quality pipelines in large, complex projects. At the highest level, the data flow of the Blackbox is a loop. Production data from projects flows into the Blackbox from an SQL database; coordinators receive an error report and make edits to the production data. The Blackbox detects these edits automatically at the next execution and removes the data issue from the list of data errors. This raises team efficiency without compromising accountability or accuracy.

2.5 Blackbox Output

Upon execution, the Blackbox generates three output files: Query Library Flags, Missingness Flags, and Missingness Summary. The Query Library Flags file serves as a catch-all for the high-level flags identified during execution, including queries added to the Query Library or issues preventing the Blackbox from fully inspecting a record. The 'Missingness Flags' file provides a detailed log of every instance where the 'required' field in the project is empty, capturing individual missing fields. The 'Missingness Summary' file aggregates data by tallying the expected versus observed completed fields across records, visits, and forms. Together, these three output files offer a comprehensive view of project status, enabling the study team to pinpoint and address issues that may otherwise go unnoticed. The files are critical in validating study statuses and instilling confidence in the data quality pipeline, particularly during mid-project statistical analyses.

These three dated output files enable the preservation of audit trails for quality improvement and regulatory purposes. Study staff receive error reports and address those within the study-prescribed timelines. Simply, when staff resolve a data issue in the project's REDCap, such as completing missing fields in a form or filing a protocol deviation, it rectifies the problem, ensuring the exclusion of such errors from future reports. Notably, direct feedback from the staff is unnecessary as the Blackbox retrieves data directly from the database.

2.6 Blackbox Requirements

The Blackbox is currently built on Python version 3.10,[2] and will not run on legacy versions of Python due to the heavy use of F-strings (introduced in Python 3.6) and case/match statements (introduced in Python 3.10). It depends on multiple packages, including *pandas*[3] for data frames and handling of large study data, *numpy*[4] for mathematics, *dateutil*[5] for simplifying operations involving dates, *sqlalchemy*[6] for connecting and writing to SQL database, and *openpyxl*[7] for processing Excel files, alongside their dependencies. For optional multithreading, the multiprocessing and logging modules were included.

3. Results

The Blackbox was first released in November 2024 for a clinical trial. The results were the expected three report files: Query Library Flags, Missingness Flags, and Missingness Summary **(Tables 1–3)**. The reports showed 1949 potential data errors. Further investigation of these errors revealed that several were due to changes made either to the protocol (in which forms and questions were marked as missing, not previously required) or to missing branching logic (which resulted in unnecessary errors). The study team addressed and resolved the errors, and missing branching logic was added. Blackbox execution was applied against the corrected dataset. A comparison

Table 1: The first three rows of the 'Missingness Flags' report.

Record ID	Location	Form	Field	Affiliated Coordinator
10	screening_arm_1	circas_rc	circas3 g	SG
10	screening_arm_1	tesic_rc	tesic8_1	SG
10	screening_arm_1	mini_upload_rc	mini_upload	SG

Table 2: The first three rows of the 'Query Library Flags' report.

Record ID	Location	Violation Type	Days Since Window Close	Affiliated Coordinator
10	screening_arm_1	Form mini_upload_rc appears to have been missed.	511	SG
10	v5_arm_1	Form blinding_evaluation_rc appears to have been missed.	513	SG
11	treatment_1_v1_arm_1	Form arisr_parent_pt appears to have been missed.	478	SG

Table 3: The first three rows of the 'Missingness Summary' report.

Record ID	Location	Form	Expected Number of Completed Fields	Actual Number of Completed Fields	Affiliated Coordinator
10	screening_arm_1	study_registration_rc	12	12	SG
10	screening_arm_1	coordinator_signoff_rc	7	7	SG
10	screening_arm_1	consent_note_rc	13	13	SG

between the two runs revealed that all data errors were resolved. The Blackbox currently executes biweekly. Results are consistently small and are quickly addressed by the respective study team.

4. Discussion

REDCap is a well-designed EDC system; however, one main limitation is the use of the Data Resolution Workflow, particularly for large-scale projects or projects with complex designs. Adoption of the Blackbox or similar auxiliary data quality processes provides a layer of quality control, resulting in confidence in a project's status, data integrity, and quick development of regulatory reports regarding missingness of data collection and Institutional Review Board (IRB) reports (such as adverse events and protocol deviations). The CDRC implemented the Blackbox in January 2025. It has yielded cleaner data and improved efficiency in developing reports for funding agencies and the IRB.

4.1 Blackbox Limitations and Next Steps

While the Blackbox has made a significant step forward in enhancing REDCap data quality at the CDRC, several opportunities for continued improvement and growth remain evident. One such feature that is not currently supported is the ability to adjust to mid-project changes where modifications to protocols alter the content of the study, such as possible changes to instruments at each visit, revised visit timelines, or the discontinuation of data collection. While executing the Blackbox with an updated format file and schedule of events, tracking between different versions is not possible. This may result in generating data error reports based on the new protocol version, which would be inaccurate. These two features will be included in future iterations of the Blackbox. Ongoing development will enable the tracking of versions over time. Additional improvements include simplifying the creation of input files and automating query assignments. Currently, information is emailed to the study team lead, who then assigns resolution activities to different personnel. This is time-consuming in large-scale research studies, in which multiple study team members in distinct roles collect the data, such as physicians and nurses. Further, determining key personnel on each visit can be challenging. Thus, sending queries to the relevant person will minimize the time for correction.

The current Blackbox process involves lengthy manual document preparation and moderate coding experience. When a planned front-end interface, along with continued feature improvements, is implemented, it will enable documents to be uploaded in standardized formats without the need for extensive coding. This will minimize errors in the format file and associated input data, continuing to improve on high-quality data collection processes using the REDCap system. Further, once these improvements are made to the current Blackbox release, the CDRC team can partner with other research groups to improve data integrity and build collaborations.

5. Conclusion

Data capture will continue to evolve and improve with advances in technology. The development of REDCap in 2004 created a significant shift in research data capture capability, allowing researchers to collect meaningful data through multiple interfaces. As with any software, proper use, development, and execution are necessary. In addition, while REDCap provides opportunities for assessing data quality, its current iteration overlooks some complexities as studies become multisite and span many years. Thus, the adoption of data quality pipelines, such as Blackbox, will aid in improving the quality and efficiency of research studies using REDCap as the method of data collection.

Acknowledgements

We would like to acknowledge the clinical staff at the Center for Depression Research and Clinical Care and our participants who have provided feedback about the design, implementation, and use of the REDCap system. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing Interests

Dr. Chin Fatt has served as an advisor for Janssen Research & Development. **Dr. Carmody** has received consultant fees from Holmusk Technologies, Inc. **Dr. Trivedi** has provided consulting services to Alkermes Inc, Axsome Therapeutics, Biogen MA Inc., Cerebral Inc., Circular Genomics Inc, Compass Pathfinder Limited, GH Research Limited, Heading Health Inc, Janssen, Legion Health Inc, Merck Sharp & Dohme Corp., Mind Medicine (MindMed) Inc, Merck Sharp & Dohme LLC, Naki Health, Ltd., Neurocrine Biosciences Inc, Noema Pharma AG, Orexo US Inc, Otsuka American Pharmaceutical Inc, Otsuka Canada Pharmaceutical Inc, Otsuka Pharmaceutical Development & Commercialization Inc, Praxis Precision Medicines Inc, SAGE Therapeutics, Sparian Biosciences Inc, Takeda Pharmaceutical Company Ltd, WebMD. He sits on the Scientific Advisory Board of Alto Neuroscience Inc, Cerebral Inc., Compass Pathfinder Limited, Heading Health, GreenLight VitalSign[6] Inc, Legion Health Inc, Merck Sharp & Dohme Corp, Orexo US Inc, Signant Health. He holds stock in Alto Neuroscience Inc, Cerebral Inc, Circular Genomics Inc, GreenLight VitalSign[6] Inc,

Legion Health Inc. Additionally, he has received editorial compensation from the American Psychiatric Association, and Oxford University Press. **Mr. Becker, Mrs. Burkhalter,** and **Mr. Dawson** report no conflicts of interest.

Author Contributions

Cherise R. Chin Fatt: Conceptualization, Methodology, Writing – original draft, Writing – review and editing, Formal analysis, Supervision.
Ryan Y. Becker: Software development, Writing – original draft, Writing – review and editing.
Lorraine Burkhalter: Writing – original draft, Writing – review and editing.
Brennan Dawson: Methodology, Software development
Thomas J. Carmody: Methodology, Writing – review and editing.
Madhukar H. Trivedi: Conceptualization, Methodology, Writing – review and editing.

References

1. **Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG.** Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–381. DOI: https://doi.org/10.1016/j.jbi.2008.08.010
2. **Phillips D.** *Python 3 object oriented programming*. Packt Publishing Ltd; 2010.
3. **McKinney W.** pandas: a foundational Python library for data analysis and statistics. *Python for high performance and scientific computing*. 2011; 14(9):1–9.
4. **Ascher D, Dubois PF, Hinsen K, Hugunin J, Oliphant T.** Numerical python. Package to speed-up arithmetic operations on arrays of numbers; 1999. Accessed 06-20-2025 at https://www.cs.mcgill.ca/~hv/articles/Numerical/numpy.pdf
5. **Niemeyer G.** dateutil: Powerful extensions to the standard Python datetime module; 2003. Available at: https://github.com/dateutil/dateutil
6. **Myers J, Copeland R, Copeland RD.** *Essential SQLAlchemy*. O'Reilly Media, Inc.; 2015.
7. **Soliev B, Odilov A., Abdurasulova Sh.** Leveraging Python for enhanced Excel functionality: A practical exploration. *Al-Farg'oniy avlodlari*. 2023;1(4):267–271. Accessed 06-20-2025 Available at: https://cyberleninka.ru/article/n/leveraging-python-for-enhanced-excel-functionality-a-practical-exploration/pdf

How to cite this article: Chin Fatt CR, Becker RY, Burkhalter L, Dawson B, Carmody TJ, Trivedi MH. Research Electronic Data Capture Using REDCap and External Data Quality Pipeline Development. *Journal of the Society for Clinical Data Management*. 2025; 5(1): 11, pp. 1–7. DOI: https://doi.org/10.47912/jscdm.431

Submitted: 23 April 2025 **Accepted:** 16 August 2025 **Published:** 03 September 2025

Journal of the Society for Clinical Data Management is a peer-reviewed open access journal published by Society for Clinical Data Management.

OPEN ACCESS