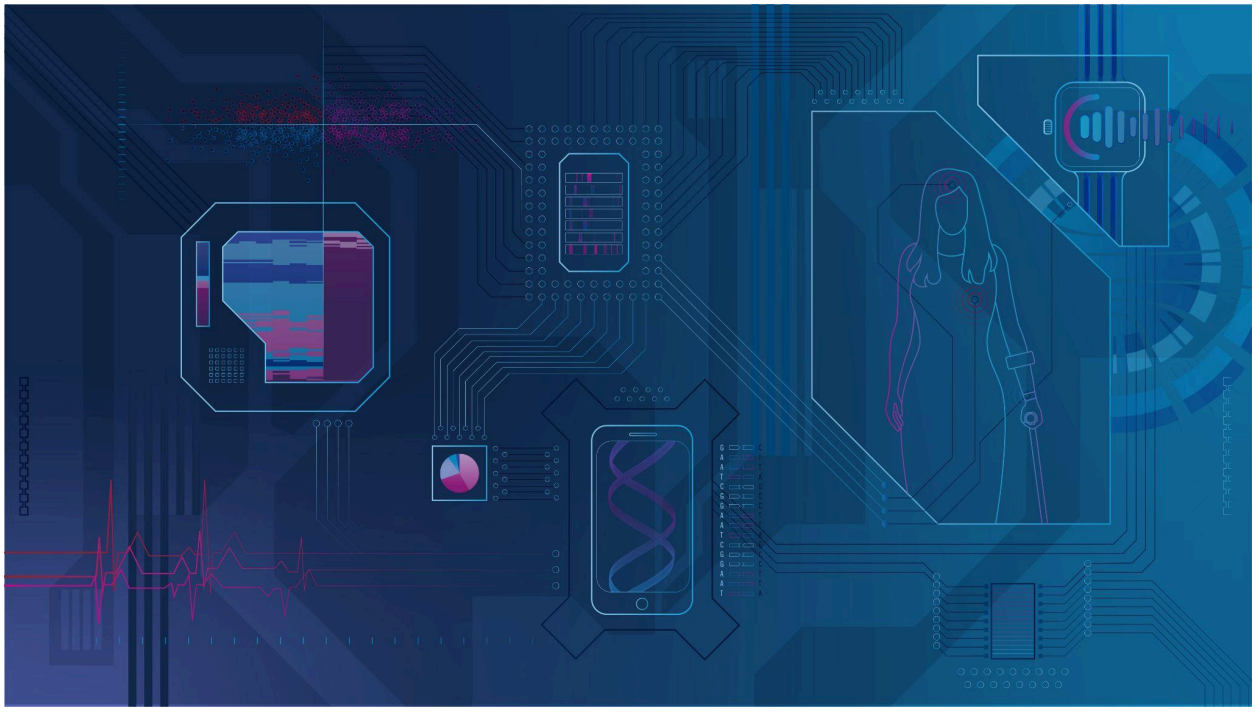
Authors: Amir Emadzadeh[1], Thea Sutton[1], Dan Otap[1], Sergio Pejenaute[2], Sara Almudéver[2], Iranzu Quintana[2], Aida Cuadrado[2], Eduardo de la Guardia[2], Sara Díaz[2], Joana Reis de Carvalho[2], Thomas Tomacci[1]
Affiliations:
1 Genentech, Inc, South San Francisco, California
2 Clinica Unviversidad de Navarra, Spain



EHR to EDC Efforts: One Sponsor's Real World Experience and Learnings

Introduction

Integrating Electronic Health Record (EHR) systems with Electronic Data Capture (EDC) systems offers a significant opportunity to streamline clinical trials, thereby reducing both time and cost burdens. In traditional clinical trials, data collection and management require duplicative efforts, where patient information must be manually transcribed from EHRs into EDC platforms. Previous research suggests that up to 70% of data is duplicative between EHR and EDC systems.[2] Additionally, transcription errors remain a significant issue, and manual data entry is often delayed due to limited resources. This inefficiency not only increases the risk of data errors and hinders the overall trial progress, it simultaneously impacts the development of new medicines.

By connecting these two systems, healthcare providers and clinical researchers can enable real-time, seamless data exchange, improving accuracy, accelerating patient recruitment, and optimizing clinical workflow. This integration can enhance trial efficiency, reduce administrative overhead, and ultimately facilitate the faster delivery of life-saving treatments to patients.

As we transition further into a digital paradigm and strive to streamline data capture efforts, it is evident that systems based on paper models are not achieving the anticipated efficiency of digital capabilities. Recognizing this, Genentech's Early Clinical Development (ECD) team launched an initiative to explore EHR-EDC integrations.

Methods

Genentech's internal software, referred to as Direct Data Connection (DDC), was developed through a careful and systematic process focused on user needs and iterative design principles. Starting with detailed requirements gathering, the system was designed for both modularity and scalability, utilizing modern programming practices such as Test Driven Development (TDD) and continuous integration.

The project's initial goal was to utilize Application Programming Interfaces (APIs) to connect EHRs and EDC systems directly. Beginning in 2019, Genentech collaborated with a leading oncology site partner to transfer EHR data, following agreed-upon mapping, directly to a sample electronic Case Report Form (eCRF) within Genentech's firewall. These data were transmitted through the site's Enterprise Data Warehouse (EDW) and included live demographics and diagnosis data, along with synthetic clinical lab, vital, and concomitant medication data for one patient participating in an active Genentech trial.

Subsequently, Genentech partnered with another U.S.-based medical center to test the interoperability of Fast Healthcare Interoperability Resources (FHIR) for the exchange of local clinical research lab data. In 2020, the teams successfully connected to the site's FHIR server to transfer synthetic safety laboratory data using Health Level Seven (HL7) FHIR standards and FHIR APIs.

In early 2020, Genentech moved forward with developing a comprehensive application for automating clinical data flow to Medidata Rave. This phase involved collaborating with additional site partners to perform parallel testing of the DDC application. During parallel testing, the DDC system and traditional data entry methods were used simultaneously in interventional oncology treatment trials, allowing for the comparison of data accuracy and consistency without affecting the primary clinical trial activities.

To transition from conducting parallel testing pilots on existing trials to utilizing live trial data, both primary and backup molecules/trials were selected for possible use. The selected molecules/studies had to meet several criteria 1) the alignment of the trial molecule's schedule with the timeline of the DDC project, 2) interest from sites in the trial molecule and the DDC solution, 3) technical capability of the site and the availability of its team, and 4) a supportive molecule team that is open to change. Efforts were implemented toward deploying the DDC technology across willing sites participating in the final selected interventional study.

Assessing Impact on Site Operations

A known barrier to the implementation of digital innovation in clinical trial ecosystems is that of scalability. The technical complexities and tailored methodologies of EHR and digital data integration exacerbate this constraint. These initiatives require significant time and substantial resources. Involving key stakeholders from all relevant parties is crucial from the very beginning of the project. Furthermore, identifying the optimal collaborative partners is important in the early stages of such efforts. Therefore, a specific site engagement strategy had to be developed to advance this initiative effectively.

Given the aforementioned reasons, the site engagement strategy for DDC expansion into live clinical trial data acquisition leveraged key existing site alliance relationships. Genentech is known for its deep site alliance relationships spanning early phase oncology (gRED Expanded Oncology Network-EON), our inclusive research network (Advancing Inclusive Research-AIR)[3], as well as our immuno-oncology (imCORE network)[4]. Leveraging these partnerships, we strategically aligned with one of our valued medical centers, the Clinica Universidad de Navarra. This center worked closely with us to successfully implement our DDC technology in one of our active interventional treatment trials across two of its research locations, each identified by separate Rave EDC site identifiers.

Clinica Universidad de Navarra enrolled its first patient in our selected trial in 2024. Upon the initial data transfer, the site remarked, “That was so easy!” Our DDC technology was deployed successfully, transferring discrete data from 214 patient visits, covering 473 submitted forms, and generating 7,331 field-level clinical data points. Site end-user feedback was strongly positive, and data entry and subsequent review by Sponsor staff were impacted positively. This achievement realized our aspirations for semi-automating the transfer of live clinical trial data from EHR to EDC.

“This was so easy!”
- Site Data Coordinator upon using the technology to push data post-implementation

Scalability

The program had achieved success in our proof-of-concept settings, and addressing scalability was identified as the next step in this digital transformation initiative. Could we deploy this solution to a substantial group of site partner types? Would our partners find it valuable? Could we realize the return on investment required to generate bespoke methodologies across our various site partners and country affiliate members?

Our plan was global, our aspiration clear, but we needed to ensure the landscape ahead was amenable to this type of effort to modernize and optimize our site data transfer ecosystems. The DDC technical staff collaborated extensively with Genentech Clinical Operations staff to conduct an updated site landscaping effort to gauge site leadership perspectives and ascertain institutional insights on the next possible steps forward.

Using a data-driven approach, the team identified the top 20% of enrollers across Genentech's early-phase oncology trials. Limiting the scope to North America, the team identified the top 27

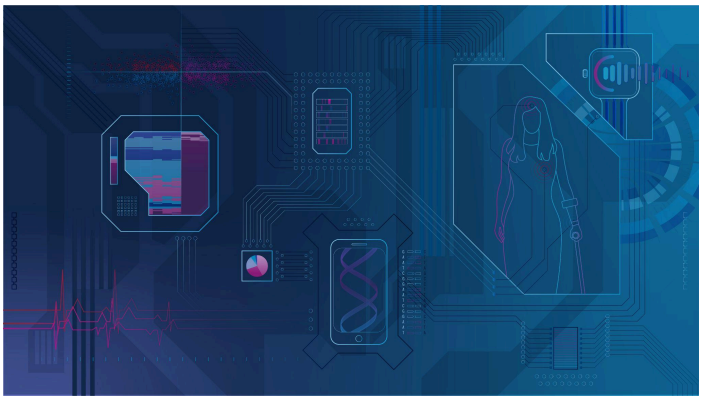
centers and began outreach in the early part of 2024. Throughout the first quarter of 2024, targeted interviews and correspondence were conducted across the centers, and a final distillation of feedback was then presented to the program sponsors for review and decision-making.

As of early 2024, our team identified key learnings from our landscape efforts, which revealed the following findings, overwhelmingly consistent across the majority of our site partners interviewed:

Feedback Analysis- Reasons Why sites want DDC

Innovation	Reduced site coordinator burden	Increased data quality
• A desire to innovate was a common characteristic across all 3 centers who would pursue DDC at this time	• 100% of sites acknowledged a net positive approach related to decreasing data coordinator time/effort related to discrete data field entry	• Decreased end user error and increased data quality confidences were highlighted as positive desired outcomes of any EHR-EDC integration effort

The majority of sites expressed interest in the EHR-EDC concept. However, various factors influenced a site's willingness and ability to deploy this type of data solution. Our landscaping analysis surfaced critical constraints in implementing DDC across the vast majority of our top site partners.

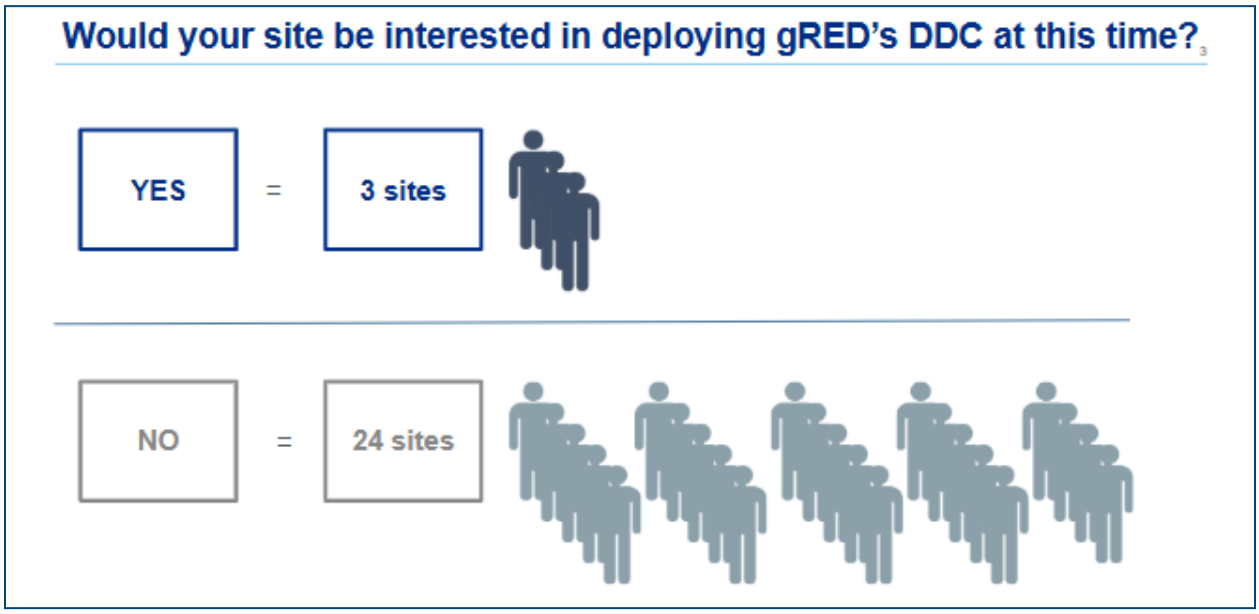


The most frequent and critical feedback was a preference for a Sponsor-agnostic solution.

Feedback Analysis- Reasons Why sites declined DDC

Unknown/Unproven ROI	IT/CTO Resource commitment concerns	DDC Program scalability concerns
• Large upfront and ongoing effort requirements cannot be justified against proposed return on investment • Cost model is crucial, outlining T/E for CTO staff and Site IT staff • Recurring IT/Staff demand on every new trial deployed	• 100% of site's raised concerns around local IT staff availability and willingness to engage on DDC • 100% of sites expressed concern related to upfront and ongoing CTO/IT staff resource allocation to maintain DDC on active studies	• 5 sites specifically listed DDC's Sponsor specificity as a driving factor to not adopt it as a solution • All sites acknowledged a preference towards a Sponsor-Agnostic solution

The landscape revealed that a substantial majority of the sites did not see a scalable and sustainable solution in a sponsor-specific EHR-EDC platform.



Current Efforts/Next Steps

Upon review of our landscaping feedback, a critical decision was made to halt the internal development of our DDC software. A large factor in this decision was the site feedback. While DDC realized its technical aspirations, scalability and site adoption issues were identified as insurmountable obstacles to effective wide-scale adoption across all site partners and ecosystems.

We determined that sites are looking for sponsor-agnostic solutions to this process of EHR-EDC, so after extensive deliberation, the decision was made to shift focus and explore external opportunities with various vendor partners specializing in the EHR to EDC solution.

With what we believe are generalizable results, we remain committed to EHR-EDC integration in clinical trials. We are unwavering in our goal to realize efficiencies in data entry timelines/quality by leveraging technology systems that surpass manual end-user data transcription. We also remain committed to being a Sponsor of Choice.

While we may not have deployed our internally developed platform across our clinical trial/site ecosystem, we remain optimistic and committed to realizing the benefits of data-transfer automation. We are encouraged by our results, and we are motivated to further explore key partner solutions on this transformative journey of digital innovation in clinical trials/healthcare.

We extend our gratitude to Clinica Universidad de Navarra, which successfully partnered with us in transferring data effectively for a live, international, multi-center interventional treatment trial. Without the commitment of their entire team, we would not have achieved this milestone. We also wish to express our deepest appreciation to the Genentech GO44010 study team and innumerable parties at other partner sites who worked with us throughout this project and contributed significantly to its success.

References:

1. Sundgren, Mats, and Nadir Ammour. "Innovations in Data Capture Transforming Trial Delivery." *Applied Clinical Trials*, vol. 30, no. 7/8, 12 Aug. 2021.
2. Sundgren, Mats, and Wai Keong Wong. "eSource Interoperability between EHR and EDC." *Applied Clinical Trials*, vol. 32, no. 7/8, 14 Aug. 2023.
3. https://www.gene.com/patients/clinical-trials/advancing-inclusive-research
4. Roche. (2016, November 15). Roche launches imCORE, a global network of cancer immunotherapy centers of excellence [Press release) Retrieved from https://assets.roche.com/imported/en/med-cor-2016-11-15-e.pdf

Special Sub-Section

Architecture and Key Features

DDC utilized a microservices architecture on a major cloud service provider platform, ensuring scalability and flexibility. This architecture allowed each DDC component to function independently, making it easier to manage and scale as needed.

- DDC Portal: A front-end application developed with the React framework and hosted on the cloud, providing a user-friendly interface for all user personas.
- DDC-sync-io API: A microservices layer built with NodeJS/TypeScript and serverless compute leveraging the GraphQL framework. These APIs can be categorized into three different modules:
    - Site Integration Module: Managed connectivity with clinical sites, handling data ingestion through both pull and push methods.
    - Data Processing Module: Managed data flows and transformations, ensuring that data is correctly formatted and validated before submission.
    - Rave Integration Module: Facilitated data interactions with the Medidata Rave system, converting data from JavaScript Object Notation (JSON) to the Extensible Markup Language (XML) formats required by Rave.
- DDC Database: A MySQL database hosted in the cloud, storing Rave metadata, raw and processed data from sites, as well as site-specific mappings. It ensured the maintenance of data integrity and consistency.

Automated data transfer supported both pull and push mechanisms and was compatible with industry standards like FHIR and HL7, and custom data structures. This flexibility allowed integration with various EHR implementations.

Data mapping is crucial as clinical sites often vary in data attribute names and structures. The DDC system enabled the mapping of site-specific attributes to standardized EDC forms, ensuring consistency and accuracy.

A user-centric design was central to DDC, tailoring interfaces for distinct user personas. This design approach allowed reviewers to verify data before submission, administrators to manage settings efficiently, and data mappers to configure mappings accurately.

The system achieved robust messaging and data management through the use of a message broker for asynchronous messaging, enhancing system resilience. Data was securely stored in object storage solutions, ensuring high durability and availability, while structured data storage in a relational database allowed efficient querying.

Ensuring security and compliance was paramount in DDC's design. By using Medidata as a Central Authentication Service (CAS) for single sign-on authentication, DDC eliminated backend credential risks. Data encryption was applied at rest and in transit, and user access was tightly controlled through integration with the Medidata Rave platform. Comprehensive audit trails were maintained to ensure regulatory compliance.

Validation of the DDC Application

A thorough validation process ensured reliability and compliance for clinical research integration. Guided by the User Requirements Specification (URS), the process included functional and non-functional testing, user acceptance testing (UAT), audit trails, compliance verification, security checks, performance assessment, and disaster recovery. This rigorous validation confirmed that the DDC application met its requirements, including application redundancy and failover, thereby enhancing clinical trial data management by securely integrating EHR with EDC systems.