

## SCDM Thoughts: RWC (Real World Clinical) and a Data Fabric for EHR to EDC

I attended SCDM (Society for Clinical Data Management) for the first time. I had never put much thought into clinical data management as someone focused on patient data from the EHR side. Even as we got into doing research studies on health system data for pharma it rapidly became coined Real World Data (RWD) as a field. As RWD it stayed separate from the clinical research side of life sciences companies. But things have been continuously evolving. In our business at Graticule we get closer and closer to clinical trials with our work every year especially given our efforts to work on the hard problems relating to improving research using health system medical data and systems. So for now I have some relatively fresh eyes for the clinical data management space as I attended my first meeting of SCDM. Here are some thoughts that I had based on the talks and conversations I had with attendees and vendors.

I feel like there is a big divide between clinical trial groups in pharma and real world data groups. In meetings with life sciences companies at SCDM and engaging chats with groups involved in EU regulatory bodies there seems to be a low understanding of what the Real World Data groups are currently able to do. It gets nuanced because the clinical trial world and especially clinical data management is very focused on submissions and the regulatory requirements for submissions. For example at SCDM, there was an hour and a half long panel attended by the broad conference attendees from regulators from the US FDA, and four other countries each providing their thoughts on innovation and directions. I've never seen such a presentation at an RWE conference. In terms of connecting on how we work in RWD, a Danish inspector asked incredulously why it was ok ethically to submit data to a regulator as a part of a trial in a submission where the patient hasn't consented. I thought about it, because in RWD we often use waivers of consent and de-identification frameworks to handle the question of whether the data is appropriately meeting privacy standards. At least in the US the FDA has provided significant guidance on the topic and has worked to make it become possible (<https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>)

But in the EU, or according to the researcher from Denmark, they have remained with a stance that it is not acceptable to have a submission where patients have not consented to being included in the study if it is prospectively collected. For example, if data is collected in a surveillance system following the use of a drug or medical device. Apparently, this remains true even when the data use is not in an RCT so the patient care is in no way impacted by collection of the long term outcomes or adverse events that the patient experiences. My point here isn't

that they should change the laws in Denmark but that there is still a divide on the perceived ethics in a clinical trial space which results in submissions to the FDA and the non-clinical trial space of RWE where most of the outputs are for publication into journals to demonstrate post-marketing factors, or for payors to understand the risk benefits for funding access to a medicine.

Yet we are converging between health system data collection tools (EHRs) and clinical trial collection tools (EDC). One big theme of the conference was how to lower site burden but still ensure effective quality control and traceability systems of the data being collected during trials. There have been good investments on the clinical side such as eSourcing. But in actually walking the floor of the conference, I learned that eSourcing is generally not getting data from an EHR so much as having the site source the data into their own system to push up to the sponsor. This does make thinking about EHR to EDC a different concept and sourcing EHR data into a sponsor system is a different concept than how eSourcing is considered today. So we are going to continue to get confused with mismatched vocabularies when talking about how to discuss how RWD and clinical development connect.

I also didn't get the impression that real world data leadership in life sciences companies attend SCDM despite RWD being on the agenda at the conference. Instead, the teams at the conference were mostly still centered on data management and while interested in real world data they see the incorporation of EHR data into trial designs as their specific model and as distinct from RWD.

One idea that came to mind was that maybe using EHR data in clinical studies used in submissions to regulators needs a new name. It isn't really EHR to EDC as many of the vendors with EDC tools, CROs, or specialized utility companies provide. We have RWD (Real World Data) which from my experience is any data outside of a clinical trial and potentially including other historical clinical trials. RWE (Real World Evidence) is the analytical output created from RWD to establish evidence that has been peer reviewed. This presents a challenge to describe the use of patient data in a clinical trial. Partially because including the patient record from an EHR into a clinical trial magically transforms the data from being real world data into becoming clinical trial data. Once it does it is subject to the processes of clinical data management and regulatory requirements that are intended to oversee clinical trial submissions to regulators. So when patient data hits a clinical trial then it stops being RWD.

Maybe the correct term for doing things bridging these areas should be RWC (Real World Clinical) or another equivalent term if there is one that would catch on, if it could be coined. Personally I like RWC because we already have an RWD and an RWE so we can have an RW(C, D, E). RWC would describe the use of patient records from electronic medical records that are not built and maintained just for the study. In order to get to a RWC concept we would



need a lot more collaboration between RWD teams which generally live in spaces outside of clinical and the clinical development teams. At Graticule, we don't have the luxury of a scale where we are separated and these all fall into a mix of epidemiology, biostatistics, data science, software engineering, compliance, and business strategy. We are happy to help groups who want to bring these two disciplines together if either side is open for establishing a consortia together.

Another impression I had at SCDM was positive from my perspective relative to how vendors offering solutions in clinical data management think about what we at Graticule are looking to do with CLEHR. I'll start with how I described CLEHR to the groups that we spoke with and proceed with where they told me that they see value in what we are building out. I mention this because your company may have similar needs and we are happy to engage to be helpful.

I described CLEHR as the following: Graticule has been focused on the interface into the health system because we come from a history of Real World Data. In the case of clinical, we are focusing on the clinical record as represented in FHIR. At Graticule we are not focused on the full study and all of the downstream operations on the data because the complexity of managing that or imposing it on sites would be detrimental to what clients we have spoken to want. We are trying to keep it as simple as possible for a health system site to adopt an interface to allow them to provide access on a limited basis to their EHR for extracting records for use in driving trial efficiency at sites. This means we won't ask sites to adopt bells and whistles like mapping tools and integrations for front-end utilities to review and copy data. Those tools may be made available but for now CLEHR doesn't fit that bill. CLEHR is focused on solving for the central trust issues as an honest broker, which are:

- The site always has control over which patients are included in the study by controlling the MRNs
- CLEHR manages identity mappings into downstream systems such as the mapping between an MRN and a researchID in an EDC with a process to do so that works efficiently for the site
- CLEHR de identifies records including free text based on the specifications of the protocol it is supporting
- Graticule signs a BAA or equivalent with the site to handle the fact that Graticule is processing identified patient data for the purpose of running a study
- CLEHR only conveys the data agreed upon in the protocol and limits it either by not requesting it or by reducing the data set prior to transmission to the downstream system
- CLEHR provides transparency back to the site and sponsor systems of what occurs
- The health system can turn off use of CLEHR at any time by shutting off access to the FHIR API



This minimal product becomes the basis for an implementation at a site that can be cut down to about an 8-hour elapsed build and a validation process that can be contained. It also can focus on a process to validate that the data being provided through FHIR is in fact the source data. We can optimize the process by focusing on the key steps. But the more important thing is that once connected to a health system, this connection can be used for many studies that accept it as the conduit for the data exchange.

One more thing needs to be introduced. CLEHR is a technology and a HIPAA-secure cloud platform, but it is more than that because getting connectivity to work involves contracts and agreements with health systems. These are difficult stumbling blocks to work through. Health systems have large backlogs in IT, big risks around HIPAA breach risk, and aren't sure of the business and reputational risk of collaborating around patient data. So while it may be a short implementation cycle per connection, it is not a short contracting cycle. We estimate it is 6 months to a year for a health system to work through the process to agree to connect to CLEHR or any system like it. The process itself is for the subject of another blog post but it explains what I'll describe later as the response of clinical data management vendor groups.

The technology and platform while non-trivial to build is not the key value proposition. The key value proposition is solving over time for the challenge of scale. In our case, solving this by limiting the scope to a capability that can be scaled effectively. So CLEHR is more a network than a technology. Rather than asking how a cell tower works most people want to know about the 'can you hear me now' problem. So CLEHR is more like a carrier with specific capabilities that can be subscribed to by either party (source and recipient) than like a piece of software. We have been fortunate to be building up connectivity in this way through large health systems to achieve close to 50M addressable patients in 2024 and the option for significant growth in 2025 now that we have established the model for connectivity, and contracting.

Now getting to the point on what we learned from the vendor groups. I'll put them into two buckets: CROs and EDC vendors. There are others but this should cover the main insights from the conference regarding EHR interfaces.

CROs: Contract Research Organizations (CROs) often outsource clinical trial projects from sponsors. They solve for the whole product of delivery of an end to end trial including the technology and management of the study. They often have been asked or have tried to innovate with EHR connectivity but it has been challenging for them in their role. They have the capacity to build the tools for improving the use of the data through mapping into EDC tools or using AI to support abstraction of content from notes but getting to the data interfaces is a pain point for them. What they have learned in trying to build out capabilities for EHR to EDC is that the cycle time on connectivity is very long. The cycle time is sufficiently long and complex that they are open to CLEHR as a connector for the EHR patient data so that they can focus on the



downstream processing as fits for the study they are conducting and for their technology roadmap. They may have other sources for doing this connectivity piece, including direct connectivity at sites and they may mix and match. But if there is a way to reduce the friction from the site into the CRO, then they are open to using an honest broker who has already gone through the process to stand up a network of health systems like we have at Graticule.

While the smaller CROs with niche capabilities are the most at risk of not having the resources for interfacing, even the larger CROs gave feedback that they have now gone through cycles to do attempts to connect and that they are open to different ways so that they can focus on their core competencies to innovate. Also this EHR connectivity is still not mainstream for them so they need to think about where it might be most applicable in their client base in terms of use cases and not distract their core teams too much until it is more mature. So they may recommend a tool like CLEHR to a sponsor who asks them 'can you do this EHR integration' or perform an evaluation for the sponsor of what portfolio of tools to use given the fact that one network may not be sufficient to solve for the sponsor needs if they are looking to connect with EHR data sets.

EDC vendors: While there are dominant tools on the market for EDC such as Medidata Rave there are also many EDC vendors offering solutions to capture data that fill the market for both price ranges, geographical diversity, flexibility of the vendor, and specialized data capture such as biomarkers. The result is that at SCDM we can meet 20 vendors who offer EDC capabilities. Some are even EDC companies who have become EHR to EDC companies for other EDC companies and operate as EDC in the middle for the sites to map data into the broader EDC tool. Some have built EHR connectors into their EDC. One I spoke with had a connector that would take a PDF of the patient record as an input and then map the information from the PDF into the EDC system. So most have encountered the question of EHR integration. Even larger EDC vendors with big software platforms struggle to get the right contracts and structures in place for EHR to EDC to work. For example Oracle has acquired many companies including Phase Forward, Cerner, and Siebel. Cerner is one of the leading EHR platforms. But to connect into the EDC systems they offer they need to be agnostic to EHR since the health systems often have Epic instead of Cerner. Thus their EHR to EDC strategy isn't entirely helped by offering Cerner and even with Cerner the connectivity from Cerner out has the same issue that they might connect to a different EDC than one owned by Oracle.

Overall the same sort of needs occurs in the EDC world of needing a middleware system with a network in place that can work in a many to many fashion between the health system EHR and the EDC system. Otherwise the contracting process is very complex. So, the EDC vendors I spoke with are also supportive of a solution where a system like CLEHR with a built-in network becomes the bridge between the EDC system and the health systems. Many have evaluated the more end-to-end EHR to EDC solutions and they face the challenge that adopting them puts

them into a position of having limited control of the technology for processing the data in the EHR output records, or to offer tools that are integrated to their EDC platform.

We are listening to this feedback and working as a follow-up with both CROs and EDC companies to improve our framework for business partnering with sites. What we are structuring is a way to handle the capacity to work with sites we have already connected to CLEHR and to bring new sites into their network as needed for studies to support their approach.

So the value proposition for Graticule CLEHR moving forward with these partnerships for the sites is to lower the number of requests for EHR to EDC technology connectivity. This allows sites to focus on the core piece of technology they are evaluating, which is the connectivity and patient identity management systems that they are opening their EHR interfaces to. By opening to a small number of tools such as CLEHR with broad connectivity to clinical trial systems, they can achieve the goals of EHR to EDC of lowering the burden on their clinical research coordinator staff and improve the quality and quantity of information provided to groups doing studies. Ultimately what it comes down to: is that in order to obtain the value from connectivity and new technologies the data fabric itself needs to be available to provide information in a secure, privacy -protecting way. Once in place, these tools can start to have the major impact predicted for lowering site burden and improving the cost structure of trials for both the site and sponsor so that research can be conducted more efficiently.

Here is a diagram of what we are envisioning:

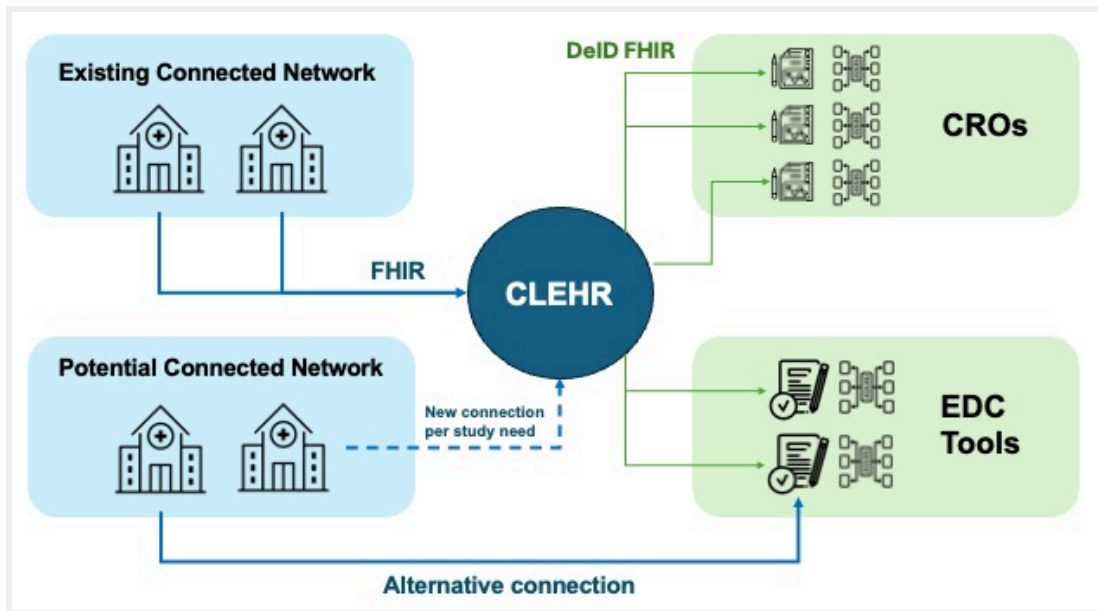


Figure 1: EHR Interface Network



In this model above, CLEHR provides the honest broker with no framework included to handle the EHR to EDC mapping workload. Each stakeholder establishes their own framework for EHR to EDC. CROs build their mapping tools and processes for handling data at an enterprise level as well as working to determine which information belongs in EDC vs. processing it downstream into other systems. EDC tools also build out their own systems for EHR to EDC mapping that are highly connected to their tools. They can then train solution providers on the EHR to EDC systems. Best of breed EHR to EDC mapping tool frameworks may be used by any of these groups including sponsors or service providers without having to factor the integration network capabilities of the EHR to EDC mapping tool or to risk the time to integrate each site if a better or more appropriate tool is available. The sponsors themselves may select to build their own frameworks for these that match their internal operating models.

While the network may begin in a current state of having lower than optimal coverage for many studies (50M patients), the network grows rapidly as new groups subscribe to the existing connectivity and add connections to support their specific needs. Once the connections are made, because there are sufficient additional programs running through the connection, it becomes useful for many studies vs. just a single study. This simplifies the clinical trial office and EHR integration team at the sites to focus on fewer IT projects to integrate with external groups to support studies incorporating EHR data into the study.

Furthermore the health systems and independent sites can invest more time into the review of the processes and establishment of effective policies or technology to govern the use of the connections. Graticule through CLEHR takes on the responsibility for the site and to protect the downstream parties to comply with patient privacy requirements and study specific requirements around data sharing by configuring CLEHR for each study to only provide the minimum necessary data and to manage identifier mappings to research systems in a secure and compliant fashion. CLEHR also achieves the appropriate levels of governance and audits to be considered compliant by regulatory groups such as FDA or EMA by establishing long term processes needed to establish source data validity in the connectivity layer.

As always, please reach out to myself or my team if you wish to discuss these topics and how they may relate to your research efforts.

Interested in learning more?

**Reach out to us at [info@graticule.life](mailto:info@graticule.life)**