



HIESourcing to Support Clinical Research with EHR Data at Scale: Pioneers Wanted

Graticule is seeking life sciences pioneers to work with our cloud infrastructure and informatics honest broker services to establish studies and systems that source data through Health Information Exchanges to populate patient data into sponsored clinical research studies. While it is early in the process for the maturity of HIESourcing the time to start structuring plans for innovation and testing pilots is now.

HIESourcing is similar to eSourcing where data is sourced from EHRs through direct interfaces but we are instead establishing connectivity through HIEs instead of individual EHRs at each site. We are calling this emerging concept HIESourcing as an extension of the existing eSourcing that is already gaining momentum. HIESourcing has the potential to bring a critical missing component to acquiring data related to consented research participants. SCALE!

HIESourcing is eSourcing at scale using Health Information Exchanges

The key to doing so is to use an informed research consent collected from patients during study enrollment that allows for automated access to, for example, the 250 million+ patients through a single national network interface such as ehealth Exchange. While this solution would be considered science fiction and impossible given the regulatory guidelines around HIEs a few years ago we believe that is about to change in 2025. So we want Graticule clients and our research community partners to have already established the appropriate protocols, consent processes, and data quality strategies in place to leverage changes we expect from ONC and other federal regulators by the beginning of 2026.

eSourcing is still held out as one of the holy grail topics for improving clinical trials. When discussing eSourcing with most clinical trial veterans they will remind you that this idea and technical approaches to solve it have been around for 30 years but with limited success at scale. The benefits of eSourcing are clear in that the original source data represents an ideal solution for SDV (Source Data Verification) to the point of potentially making SDV as we know it extinct. Most clinical data managers will attest that if the source data itself from electronic health records was made available throughout the study then the process of SDV with coordination through calls back and forth with sites to check the original EHR record is accurately reflected in the EDC becomes unnecessary. This is because the source is available to the sponsor to verify when a potential issue is identified. SDV is primarily making sure that the source, the EHR, matches the research record in the EDC tool or some other system used in a study.



Unfortunately eSourcing has always been a difficult to execute project. It continues to be a challenge because of the complexity of quickly, cost effectively, and securely establishing trusted technical links between health system EHRs and sponsor clinical trial data infrastructure. With studies being on tight time frames for execution there never seems to be a good time to invest in build out of networks.

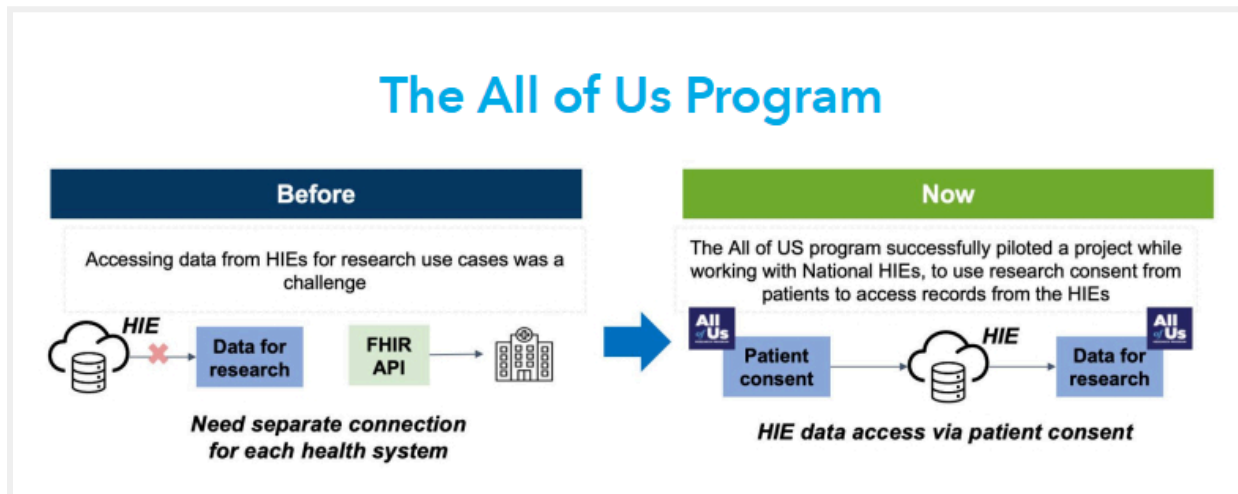
At Graticule we have been tackling the eSourcing problem with a new approach that leverages the 21st Century Cures Act that requires EHRs to offer FHIR interfaces and a mandate to comply with USCDI V1 from the ONC (Office of the National Coordinator). Mandatory FHIR interfaces are a giant leap forward primarily because health systems can now deploy interfaces to trusted parties like Graticule to share patient data on appropriate patients through our cloud platform CLEHR. The big change that may be the most important to health systems is that the API costs from EHR vendors are \$0 because these APIs come as a part of their core license for the EHR product.

But, as we have seen in our deployments at major health systems, it is still not trivial to go through the process of integrating with a health system in order to offer the plumbing to connect to clinical research sites. For each health system and pharma looking to connect for one or more studies it can be a long process including CORL audits, business justification at the health system, contracting, and IT integration team projects. It can take 3-8 months to go through the contracting and planning cycle to get a new site live for eSourcing due to the complexity of the process. After 3 years Graticule now has established links that can scale to over 50M lives but it is still far from the full set of patients in the US which is over 300M. We estimate that it will take another 3-5 years to build-up a network to have the coverage most sponsors are seeking on day one to be able to know when they initiate a trial that all of their selected sites will be online by the start-up phase for each site. Furthermore this process still risks limiting access to trials and available patients who can participate. Many community sites which care for patients not represented in research are harder to network to because they are not part of large AMC networks that normally can contract through their clinical trial office.

At Graticule we have always been interested in the potential to source data from a broader pool of patients than the health system sites running trials. Furthermore we have been looking into ways to help distributed studies which don't rely on large AMCs (Academic Medical Centers) to access patient records. One solution is to have patients retrieve the records themselves through their logins via portals like MyChart. Unfortunately that leaves a lot of uncertainty on whether records can be expected on each patient so it is generally only applicable for observational registries and not interventional studies like RCTs (Randomized Controlled Trials). Even with observational registries it can be a difficult approach for both sponsors and patients given that most patients don't know how to use their logins for EHR access to their records or remember all of the health systems they have interacted with.

But HIEs operate behind the scenes and already offer information exchange for the purpose of clinical care. Patients are not the operators of HIEs so providers with experience can manage the processes within HIEs. The connections to HIEs from health systems are already complete in order to meet the requirements for information exchange between health systems to comply with ONC guidelines. With a few changes suddenly data can become available from HIEs into research studies. That primary change is how the HIE is allowed to respond to a valid research consent from the patient. With a model where the consent provided by the patient can be conveyed electronically along with the request for access to the patient record then a research system can access the HIE to act on the patient's behalf to retrieve their EHR data.

Who could authorize a change allowing this? The group that governs the HIEs, ONC and possibly also Congress if needed! Is this going to happen soon? We believe it is imminent in 2025 pending the results of an initial pilot project focused on solving this problem for a high profile NIH biobank. While details are still emerging on when guidelines will be expanding beyond an initial pilot and thus available to industry there is already a first instance of doing this consent driven HIESourcing. A pilot is already underway with data flowing and being analyzed now.



In 2022 the NIH sought to improve the way that EHR data was associated with the All of Us research program. All of Us is a cohort where patients agree to share their medical records along with genetic samples collected at participating sites in order to connect genetics and phenotypes. Their primary collection method was to obtain feeds from each site about the patient in OMOP data feeds that were sent to a central collection system at Vanderbilt. The challenge with such a system is that it only incorporated EHR records at the site enrolling the patient. Furthermore, using batch data files from data warehouses often limits data that can be shared because the warehouses each have different capacities regarding the scope of data that

is provided to the warehouse and mapping frameworks to translate data into OMOP. The actual connectivity with flat files ends-up being many bespoke processes including challenging governance to ensure common adherence to mapping concepts into OMOP where there is ambiguity on how to standardize data. Another big issue for using the site data warehouse approach is that patients and patient records at AMC sites do not effectively represent the whole US population. Many patients live in rural areas or are provided care by community health systems that aren't a part of an AMC. So the records in All of Us, or any research system pulling data from an enrolling site health system is limited. But the actual informed consent for All of Us that the patient agreed to isn't limited to the site. While there are complexities with the evolution of consent forms over time the general core consent for All of Us is for the patient to share all of their electronic health records for the purposes of the research biobank. NIH reached out with this RFI in 2022 describing the interest in the problem.

(Source: <https://grants.nih.gov/grants/guide/notice-files/NOT-PM-22-004.html>). Note the following text from the 2022 RFI: "Alternative Approaches to Collecting Electronic Health Record Data for All of Us")

All of Us has explored different approaches to overcome this gap. One approach being utilized is through participant-mediated mechanisms. In this mode, participants provide All of Us with access to their EHR data through patient portals and Fast Healthcare Interoperability Resources (FHIR) based connections. Currently, this approach is available to a subset of participants. Of these participants about 8,000 (75%) contribute EHR data from one healthcare provider, while another 1,400 contribute EHR data from two providers.

An emerging area of focus for sharing EHR data is through Health Information Networks (HIN) and Health Information Exchanges (HIE). These organizations support sharing EHR data for purposes such as patient care, public health surveillance, emergency services, and patient requests. They also support sharing for additional purposes reliant upon authorizations to disclose protected health information (HIPAA authorization). All of Us obtains a HIPAA authorization from participants who agree to share their EHR data. Consequently, there is no additional participant burden to seek EHR data through an HIN or HIE. Adding EHR data for All of Us participants from HINs and HIEs would partially fill data gaps, improving the completeness of EHR data available to researchers. Therefore, we are seeking information on using HINs and HIEs as another source of EHR data"

If we now fast forward to the end of 2024 the project was funded in 2023 and proceeded through implementation through the All of Us [CLAD](#) (Center for Linkage and Acquisition of Data) now [operated at the University of North Carolina](#).



At this stage we don't have public reports but we expect to hear from the team and Graticule may work with the team in Q1 2025 to report out whether it is operational where the governance has been worked out and the technology is in place. We understand that ONC has been supportive of the All of Us consent being sufficient to acquire records from large national HIEs such as eHealth Exchange and that data may be already flowing into an evaluation framework from more than one HIE in an project to determine if the data collected through FHIR from large HIEs and HINs are as good as or sufficient to supplant and/or supplement the direct EHR OMOP feeds used in the national biorepository. So the technical components were sufficiently funded and significant work has been put in place in 2023 and 2024 to enable these connections for communicating the consent as a component of required privacy protections. If successful the project will transition into a production mode with authorization for full data use in 2025 vs. the evaluation and open the precedent and approach for all research projects with the same informed consent. While the wheels of change turn slowly this is a next big opportunity for study teams looking to eSource data at a national level. We are monitoring progress through opening dialogs with the teams involved to determine how to translate this work effectively into industry projects where similar business issues arise around tracking consented patients through EHRs.

Certain critical path applications such as cell therapy long term follow-up registries, rare disease registries with genetic biobanks, and post marketing safety studies would likely be the first applications to make primary use of this broad research consent and retrieval model. But a longer term vision would be to enable this for the wide variety of study designs considered to be clinical trials including observational studies, phase 1-3 randomized clinical trials, and decentralized trials to significantly reduce the burden on sponsors for funding manual data collection.

What to do now?

Now is the time to invest in identifying which research initiatives are appropriate for this and to establish convening among internal and peer groups. It is also time to establish appropriate engagement with the key stakeholders such as ONC and major health exchanges to initiate the first non-government pilot study applications. We are working to help our client teams to learn the details of what has been done to date and to understand the constraints of the model relative to existing approaches including how to combine existing eSourcing approaches with an HIEsourcing approach where logical.

Now is also time to review clinical research studies where use of this type of resource is anticipated. One risk is that protocols that don't anticipate these types of resources will miss out on the opportunity when it emerges. So factoring future proofing active and planned protocols and consent can now include sufficient foresight to leverage the potential of these resources if



only to collect and manage consent data that is compatible with how the All of Us project operated.

Graticule, sponsors, and other industry partners can accelerate this process by working together to bring life science companies together to communicate the benefits and their side to a commitment to a framework with strong protections for privacy that ONC can incorporate into guidelines to direct whether HIEs can support HIESourcing for industry.

Also many groups have still been cautious to begin eSourcing to date and this means that readiness for a large volume of EHR data is low. If your organization has not begun eSourcing then working at a single health system or a small portfolio of health systems today to interface through FHIR through a service like CLEHR will also enhance your company's maturity at handling EHR data in clinical data management so that those systems can be applied to the big data coming on the horizon. At Graticule we can work with health systems active with interfaces today to run pilot projects or to implement capabilities for multiple studies as a mechanism to increase sponsor maturity. We can also work to help your team and ours fill the gaps with a combination of services and tools such as the deID framework in CLEHR, data visualization frameworks for data quality, and mapping systems into EDC or RBQM systems.

As with many visionary programs this will be a journey and may hit some roadblocks along the way. But are gearing up for action and seeking life science visionaries who see the potential and can invest toward making it a new reality.

Contact us at ops@graticule.life if you are interested in being a pioneer, and/or you would like to hear future updates about HIESourcing from the Graticule team.

Interested in learning more?

Reach out to us at info@graticule.life