Use of an Electronic Health Record to Optimize Site Performance in Randomized Clinical Trials

Yvette Henry, Valerie Harkins, Ashley Ferrari, Peter B Berger*
Cardiovascular Center for Clinical Research, Geisinger Heart Institute, Geisinger Medical Center, Danville, PA, USA

Abstract

Electronic Health Records (EHRs) are widely recognized for their potential to revolutionize healthcare delivery and research. Less established are the transformative ways in which EHRs may influence the performance of clinical trials. EHR systems provide a valuable mechanism for assessing potential research trial populations, recruiting patients into trials, and enhancing trial efficiency, cost-effectiveness, quality, and accuracy [1].

Geisinger Health System has been using an EHR since the late 1990s. All care delivered at Geisinger-inpatient and outpatient care at any Geisinger hospital or community practice site - is recorded within the EHR. We review the use of EHRs to recruit for 5 large randomized clinical trials at Geisinger, reporting the power and pitfalls associated with its use. The utilization of the EHR can transform how clinical trials are performed, but changes within the industry will be needed to fully realize the benefits.

Keywords: Electronic health record; Clinical trials; Recruitment; Randomized; Research; Chronic disease


Introduction

Electronic health records (EHRs) are widely recognized for their potential to revolutionize healthcare delivery and research. Less established are the transformative ways in which EHRs may influence the performance of clinical trials. EHR systems provide a valuable mechanism for assessing potential research trial populations, recruiting patients into trials, and enhancing trial efficiency, cost-effectiveness, quality, and accuracy [1].

Researchers have reported using EHR systems and other sources of electronic health information for alerting clinicians or study staff about potential trial subjects [2-9]. They have also established the feasibility of using electronic health information to identify trial-eligible patients [10-19] and to recruit eligible patients into trials [20-25]. The strategic use of electronic health information for identifying eligible patients with qualifying conditions to enroll into prospective clinical trials has time-savings benefits [18-21] and may be more impactful than previously reported [25-26].

Geisinger Health System is a fully-integrated, non-profit health system that serves more than 2.6 million residents in 44 counties in Pennsylvania, including several designated as rural/underserved. Geisinger owns 5 tertiary hospitals, 4 community hospitals, and >40 clinics throughout its service area. The same EHR platform is used by clinicians at every Geisinger facility for all inpatient and outpatient care.

In 2006, Geisinger created a comprehensive enterprise-level data warehouse to be the single source for its clinical, financial, operational, and research needs by housing cleansed and normalized data in a common repository, the Clinical Decision Intelligence System (CDIS). The warehouse, which contains data from over 2,500,000 patients, is updated every 24 hours with feeds from multiple data sources including the EHR, laboratory medicine, tumor registry, financial decision support, claims, patient satisfaction surveys, and high-use third-party reference datasets.

We report on the use of the EHR and other health information databases for optimizing enrollment into clinical trials at Geisinger, reviewing the power and pitfalls associated with its use.

Methods

The Cardiovascular Center for Clinical Research at Geisinger uses its EHR to systematically identify patients who appear eligible for several randomized clinical trials, and has competitively enrolled in these trials as a result.

Five recently-performed clinical trials illustrate the effectiveness of using the EHR to recruit for randomized, prospective clinical trials. Table 1 describes the key characteristics of the 5 studies: STABILITY, AWARD, The Light Study, Odyssey Alternative, and ACCELERATE [27-31].

Our recruitment approach was similar for all 5 studies. We included all of the inclusion and exclusion parameters from each study protocol that could be programmatically identified in the EHR in a search of potentially eligible patients using targeted queries that specified International Classification of Disease 9 (ICD-9) codes for each relevant diagnosis, vital signs, and allowed/disallowed medications. When Current Procedural Terminology (CPT) codes for pertinent procedures were part of the criteria, these were identified by cross-referencing medical record numbers in the EHR with the billing database within CDIS. These criteria included, for example, qualifying diagnosis or condition; qualifying vital signs; presence or absence of...
relevant co-morbidities; medications; and demographic factors such as age and sex. In initial study phases, these carefully programmed reports, or data pulls, from the EHR automated much of the traditional manual work of patient identification and chart review, saving time and money. As programmatically-generated results were screened and refined, manual reviews were reserved for non-programmable or acute criteria at later steps in the process. Searching many different areas of a patient's electronic medical record for the identification of inclusion and exclusion criteria is necessary to take maximal advantage of the EHR. To do so most effectively, the research team works with clinicians knowledgeable about both the disease of interest and research, and with experienced data analysts who are particularly skilled at the use of the EHR for electronic searches.

An understanding of the way physicians enter data in the EHR is critical to accurately programming the data pull. For example, a diagnosis listed for a single encounter may have been included solely because the physician was investigating whether the condition was actually present; it may be necessary to require that patients have had 2 encounters with a given diagnosis before considering them to have that diagnosis in order to more accurately identify the target population programmatically.

After the data pull, letters were mailed to patients identified as potentially eligible for the study. Geisinger's Institutional Review Board (IRB) reviewed and approved all such invitational letters, requiring that they be sent from providers who have been involved in some aspect of the patient's care, whether the provider was a physician, nurse, or PharmD. The IRB also allows the director of a Geisinger clinic to sign the letters of invitation in place of individual Geisinger providers within that clinic. Mailings are kept to fewer than 500 letters at a time, and quality checks are performed to reduce the risk and impact of errors.

An important decision the study team must make for each trial is whether to use an opt-in or opt-out recruitment strategy. In the opt-in method, the patient must contact Geisinger after receiving a letter to be considered for participation. In the opt-out method, an individual could elect not to be contacted again for the study by mailing back an addressed/stamped postcard or calling the study hotline phone number, both included with the invitation letter. If a patient did not respond by one of these methods within 10 business days of mailing the letter, the study staff was permitted to call them to determine their interest and invite them to participate in the study. The choice of recruitment strategy is based on many factors, including the number of potentially eligible patients identified via the data pull, the anticipated appeal of the study to patients, the sponsor's expected enrollment closure date, and staff availability.

A multi-clinic, multi-study structured query language (SQL)-based database was used for all 5 studies. This allowed Cardiovascular Center for Clinical Research staff at multiple Geisinger locations (some separated by 80 miles or more) to track the disposition of potentially eligible study patients identified via the EHR. Research staff were able to record and update information about patients' interest in the study, their eligibility in terms of each enrollment criterion, and the status of steps/level of completion within a patient flow (i.e., ready for manual EHR screening, EHR-screened, ready for scheduling, etc.). Patient-specific status notes were added when relevant.

For studies whose inclusion/exclusion criteria required less medical knowledge, we used Geisinger's on-site Survey Research Unit. Trained interviewers (who are not medically sophisticated) used Geisinger IRB-approved phone scripts to assess patients' interest and, for some studies,
to collect additional preliminary screening information. Following the telephone contact, the Windows-based Computer Assisted Telephone Interview (Win-CATI) system sent each patient's disposition (i.e., interested or not interested) directly to the multi-study database. Patients identified as “interested” then moved into the next phase of the recruitment screening process. As eligible patients were identified for further screening, the applicable screening visits, clinic consents, and study enrollments proceeded as they would in any traditional trial. Due to the large volume of patients identified by the analysis of our EHR and, in some studies, short enrollment windows, not all interested or follow-up patients were able to be called.

Results

For the purposes of Geisinger’s study recruitment totals, the 2 to 5 Geisinger clinic locations were considered to be a single site since the methodologies described were applied and managed centrally throughout the EHR recruiting process and only the locations of the patients’ study visits differed. Often, for regulatory purposes, the sponsor considered distinct Geisinger locations to be separate sites in reporting their recruitment totals.

In the STABILITY study to test whether darapladib can safely lower the chances of having a cardiovascular event (such as a heart attack or stroke) in people with coronary heart disease [27], 5,300 Geisinger patients were identified as potentially meeting broad inclusion criteria (Figure 1). The opt-out method was used for this study. A total of 1,687 invitational letters were sent to patients. Initially, 380 patients (23%) responded to the letter. After learning more about the study, 208 patients (55%) remained interested. There were 1,307 patients who did not respond to the letter;
many of whom were contacted by phone. Of those 1,307 patients, 379 (29%) expressed interest in the study after being contacted by research staff. A total of 148 patients were ultimately screened, and 101 patients were enrolled within 6 months. Total enrollment was higher for Geisinger than any of the 165 other participating U.S. institutions, and higher than 41 of 45 participating countries. Overall, study enrollment was completed so quickly that it was closed before 92 additional patients who had expressed interest in the study could be contacted.

In the AWARD study, LY2189265 was compared with placebo and with exenatide to determine if it was effective and safe in reducing HbA1c in patients with Type 2 diabetes who were taking metformin and pioglitazone [28]. There were 2,555 patients who met the study’s broad inclusion criteria and were sent letters inviting them to participate in the study (Figure 2). An opt-out approach was used for this study. Initially, 351 of the 2,555 patients (14%) responded to the letter. After learning more about the study, 131 of the 351 (37%) spontaneous callers remained interested. Of the 2,204 patients who did not respond to the mailing, 384 (17%) expressed interest in the study after being contacted by research staff. After screening visits and a lead-in period on study drug, the Cardiovascular Center for Clinical Research enrolled a total

![AWARD Recruitment Strategy and Numbers](figure2.png)

Figure 2: Award-1 Recruitment Strategy and Numbers.
of 36 patients within 8 months, which included 3 months during which the sponsor temporarily stopped enrollment. Geisinger’s enrollment was the highest among the 89 participating U.S. sites.

In the Light study, naltrexone SR/bupropion was compared with placebo to determine whether the drugs had any impact on the frequency of major adverse cardiovascular events (MACE), including cardiovascular death, non-fatal myocardial infarction, and non-fatal stroke, in overweight and obese subjects with diabetes and/or other cardiovascular risk factors [29]. The broad eligibility criteria identified 9,447 patients; 8,310 patients were sent invitational letters (Figure 3). Because of the large number of patients identified, this study was submitted to Geisinger’s IRB using the opt-in method, and 7,661 patients who did not respond to the invitational letters were not further contacted. Of the 649 patients (8%) who responded to the letter, 478 (74%) remained interested after learning more about the study. Despite the low percentage of potentially eligible patients who responded after a single letter of invitation (8%), 116 patients were randomized in 4 months to treatment out of the 239 patients who signed informed consent. Almost half of those who provided consent did not pass further inclusion/exclusion screening.

Figure 3: The Light Study Recruitment Strategy and Numbers.
In the Odyssey Alternative study, alirocumab was studied in patients with primary hypercholesterolemia who were intolerant to statins and had at least a moderately high risk of CV events [30]. An opt-out method was used for this study; 3,939 patients met broad eligibility criteria, 3,859 of whom were mailed invitational letters (Figure 4). Initially, 247 (6%) patients responded to the letter. After learning more about the study, 141 (57%) spontaneous callers remained interested. Of the 3,551 patients who did not respond to the letter, 479 were called by research staff before enrollment closed. Of those 479 patients, 61 (13%) expressed interest in the study. In total, 14 patients were screened and consented; 12 of these patients were enrolled and randomized. All of this was accomplished within 2 months, at which time the sponsor ended enrollment.

In the ACCELERATE study, the efficacy and safety of evacetrapib was evaluated in patients with vascular disease at high risk for vascular outcomes [31]. A search of the EHR identified 6,802 patients who met broad eligibility criteria, 6,607 of whom were mailed invitational letters (Figure 5). An opt-out method was used for this study. Initially, 871 patients (13%) responded with interest to the letter and after learning more about the study, 548 (63%) remained interested. Calls were made...
to 315 patients who had not responded to the letter; 36 (11%) of these expressed interest in participating. The Geisinger Cardiovascular Center for Clinical Research screened 142 patients, 116 of whom were enrolled and randomized. All recruitment activity was performed within 8 months. Geisinger’s enrollment totals placed them 4th out of all 229 U.S. sites and 10th out of 647 sites worldwide.

Discussion

The most important finding from this analysis is that patient enrollment can be facilitated using an EHR to identify potentially eligible patients more rapidly than typically achieved by traditional means. For the 4 studies in which an opt-out method was used, the proportion of patients responding with interest who remained interested after learning more about the study was 2-6 times higher than the proportion of non-responders who were interested after being contacted by study staff. Such an approach should be financially beneficial to both sponsors and sites, enhancing their return on investment.

The impact of Geisinger’s EHR-based approach to enrollment was analyzed for these 5 double blind, randomized, placebo-controlled clinical mega-trials, all of which involved chronically-ill patients, who
are more likely to have multiple co-morbid illnesses. This underscores a particular strength of EHRs in research: the added value it provides in enrolling patients with chronic conditions. The EHR is of much less value for enrolling patients in trials of acute episodic illnesses, such as an acute coronary syndrome or decompensated heart failure requiring admission to the hospital. This is true for several reasons. First, chronic illness will generally involve more documentation of the illness in patients' EHR; this makes it more likely that each patient's record will contain multiple fields pertaining to the target condition, and thus making it easier to programmatically identify critical information such as inclusion/exclusion criteria. Second, chronic conditions affecting large populations allow narrow criteria to be used in designing database queries and filtering criteria (i.e., higher specificity that still allows for a large number of eligible patients to be identified). Studies involving diabetes and cardiac disease, for example, are particularly compatible with programmatic identification of possible patients from EHR data. In contrast, patients with an acute coronary syndrome (e.g., ST-segment elevation infarction) generally access the health system through a very narrow corridor—the cardiac catheterization laboratory or the emergency department—and therefore the identification and screening of potentially eligible patients is less influenced by the ability to perform an electronic search of a very large number of patients using an EHR.

**EHR recruitment—“friendly” inclusion/exclusion criteria**

Most trials have some criteria that are easy to identify in the EHR, such as age and gender which can be considered EHR—“friendly.” Other criteria are more difficult to identify using an EHR. Imaging results, for example, are often entered as progress notes in Geisinger's EHR rather than as discrete data fields, making specific image inclusion/exclusion criteria difficult to identify programmatically. Also, the question of which field of the EHR is the most effective to capture a given criterion is essential and less obvious than might otherwise be predicted. For example, a target diagnosis may be identified using billing codes, encounter diagnosis, problem lists, laboratory values, medications administered, or primary or secondary discharge diagnoses in the event of hospitalization, among others. None of these methods will exactly coincide with another, so depending on whether sensitivity or specificity is more valued in the identification of a potential pool of patients, one or the other method, or a combination of approaches, may be preferred. Several groups have explored the feasibility of natural language processing (NLP) to identify free-text eligibility criteria [32-34] and, while labor intensive, these tools show promise for systematically identifying eligibility criteria that are less simply programmed and would otherwise require manual record review.

For the most part, the more EHR-friendly the protocol inclusion/exclusion criteria are, the less need there is for manual review of the records of potentially eligible patients. The protocol with the least programmable criteria of the 5 trials reviewed was the Odyssey Alternative study. The “statin intolerance” inclusion criteria for this study were defined by maximum daily doses and reasons for intolerance, both difficult to programmatically identify. Although a patient's medications are easily identified using the EHR, the specific doses, frequency of administration, and timing of adverse reactions are difficult to program, so a manual review of the EHR to confirm statin intolerance was nonetheless required. Most patients whose physicians believed them to be intolerant of statins were found not to meet the strict protocol definition of intolerance to statins.

The completeness of the cohort depends on the programmability of inclusion/exclusion criteria. Sponsor consideration of the programmability of criteria during protocol development will enhance the use of EHRs for recruitment.

**Sensitivity vs. specificity—how wide to cast the net**

In general, electronic identification of eligible patients has been shown to have high sensitivity and specificity [20-21] compared with manual screening, but there may be instances where sensitivity needs to be sacrificed for specificity, or vice versa. For trials in which there are a great number of eligible patients, more specific programming methods (i.e., a “narrow net” to identify fewer false positives while missing some true positives) might be most beneficial; many eligible patients will still be identified, and the time and expense associated with manually screening, interviewing, and examining patients who will ultimately be found not to be eligible is spared. For trials in which few eligible patients exist to begin with, more sensitive but less specific methods (i.e., a “wider net” to capture all true positives while also getting some false positives) might be more beneficial. Other factors that influence how sensitive and specific the programming of the EHR data pull ought to be include the sponsor's projected enrollment window and the study budget.

**EHR recruitment budgeting**

One of the challenges of our EHR recruiting efforts has been the need to convince sponsors to reimburse these non-traditional recruitment methods. As noted recently by Kramer and Schulman [35], sponsors continue to function on yesterday's cost paradigm, an outdated business model that has not evolved with technology despite international evidence of the time and cost savings that result from using electronic health information to identify eligible patients [18-21]. Our model combines recruiting and pre-screening/screening into a single, largely automated process. The costs for these activities (i.e., programming, letters) fall outside the typical budget paradigm but the return on investment was high. Negotiations with sponsors to justify this approach often resulted in delays that contributed to narrowing the enrollment window for us to maximize enrollment. Sponsors must recognize the potential return on investment of these newer recruitment methods and, more importantly, empower those negotiating the contracts for them to make exceptions.

Geisinger's successful use of the EHR for subject recruitment into the STABILITY, AWARD-1, The Light Study, Odyssey Alternative, and ACCELERATE studies illustrates both the effectiveness of using the EHR for recruitment in randomized clinical trials and the opportunity to conduct additional studies making use of similar enrollment methods and patient populations. Particularly in the case of studies examining chronic conditions such as diabetes and cardiovascular disease—with broad, programmable inclusion criteria, the EHR can assist with identifying, recruiting, and enrolling larger numbers of study participants in shorter periods than is possible without the use of these novel techniques and an EHR.

The power of an innovative approach, EHR data, in-house Survey Research Unit call center support, experienced data analysts, and a multifaceted recruitment methodology, when combined with a large, stable patient population, enables Geisinger to foster highly effective approaches aimed at enrolling patients into randomized, prospective clinical trials. The benefits of these methods are expected to increase greatly as complex protocols are simplified and replaced with large simple trials, with more EHR-friendly inclusion and exclusion criteria.

**Registries**

These methods can also be employed for enrollment into registries.
We recently enrolled 429 patients in just 2 months in the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation, a multicenter, prospective registry intended to identify treatment patterns associated with atrial fibrillation [36]. Due to the minimal-risk nature of most registries, very high enrollment rates can be achieved through these methods.

Use of the EHR for feasibility analyses

Physician investigators are known to overestimate the number of patients who will be eligible for enrollment into a planned clinical trial. Having access to electronic health information can be useful in estimating an accurate number of potentially eligible patients [37]. The Geisinger Cardiovascular Center for Clinical Research uses the EHR when completing sponsors’ feasibility questionnaires to identify the true number of patients who appear to be eligible for the trial, as well as the Geisinger clinics where the majority of these patients are seen [38]. As with recruitment, the quality of this estimate is dependent on the ability to programmatically identify the inclusion and exclusion criteria within the EHR. This methodology may be used during the development of protocols to optimize recruitment times by determining which criteria have the biggest impact on potentially eligible patient numbers [10-19].

Conclusions

Geisinger Cardiovascular Center for Clinical Research’s experience demonstrates the ability of an EHR to effectively and efficiently identify large numbers of patients who are likely eligible to enroll in a clinical trial. Spontaneous responders are more likely to be interested than those who are contacted after not responding to a recruitment letter. These methods can dramatically reduce the expense of clinical trials and other types of research studies. An interested and generous population and innovative institutional leaders are also required to take maximal advantage of the opportunities afforded by an EHR and data warehouse.

Harnessing the full potential of the EHR requires more cleverly designed clinical trials. Incorporation of more EHR-friendly inclusion/exclusion criteria will facilitate the use of the EHR for recruiting.

We hope that delays caused by refusal on the part of sponsors and contract research organizations (CROs) to allow budgeting of these exclusion criteria will facilitate the use of the EHR for recruiting. This methodology may be used during the development of protocols to optimize recruitment times by determining which criteria have the biggest impact on potentially eligible patient numbers [10-19].

References


28. A Randomized, Placebo-Controlled Comparison of the Effects of Two Doses of LY2189265 or Exenatide on Glycemic Control in Patients With Type 2 Diabetes on Stable Doses of Metformin and Pioglitazone (AWARD-1: Assessment of Weekly Administration of LY2189265 in Diabetes-1). ClinicalTrials.gov.


30. A Randomized, Double-Blind, Double-Dummy, Active-Controlled Study to Evaluate the Efficacy and Safety of REGN727/SAR236553 in Patients With Primary Hypercholesterolemia Who Are Intolerant to Statins, Odyssey Alternative. ClinicalTrials.gov.


