Virtual Research: What It Is and What It's Doing in the Real-World Setting

David Thompson, PhD, Syneos Health, Boston, MA, USA

Confusion abounds as to what virtual research is, what it should be called, and whether there are distinct types of it in the real world.

Introduction

Virtual approaches to clinical research leverage digital technologies to relieve study sites of many, if not all, responsibilities of the research process—from identifying potential study subjects to screening them for eligibility to obtaining their consent for enrollment to entering their study data.

Such approaches have the potential to unleash the power of the patient by bringing the research process to patients versus requiring patients to bring themselves to the research process. Doing so makes sense, as statistics suggest that less than 5% of the population ever participate in clinical research (even though the vast majority report being willing to do so), and study location ranks second only to receiving placebo among the most disliked aspects of clinical trial participation.[1] There are also cost savings at stake, as reductions in site involvement and investigator burden associated with virtual approaches fuel expectations for corresponding reductions in the costs of clinical research. Finally, during the COVID-19 pandemic, methods for maintaining trial continuity while reducing face-to-face interactions between patients and trial personnel are being embraced enthusiastically. It is no wonder, then, that biopharmaceutical companies are actively seeking opportunities for "going virtual" in their clinical development programs.

> "During the COVID-19 pandemic, methods for maintaining trial continuity while reducing face-to-face interactions between patients and trial personnel are being embraced enthusiastically."

But their enthusiasm is tempered by a lack of understanding of virtual approaches, inadequate experience with digital tools for data capture, and, most importantly, the risk of things going wrong in their all-important phase II-III clinical trials. A recent survey asked manufacturers to list the biggest challenges they are facing in adopting virtual clinical trials.[2] While 17% said they simply "did not know how to start," 23% cited "perceived regulatory risk" and 38% pointed to "risk associated with novel technology" as the problem. These concerns, along with the naturally simpatico relationship between digital technologies and real-world measures, have led to a disproportionate growth in the use of virtual approaches in the real-world setting as opposed to randomized controlled trials. Nonetheless, it is still the case that confusion abounds as to what virtual research is, what it should be called, and whether there are distinct types of it in the real world. The objective of this paper is to bring clarity to these issues.

Virtual Research: What Are We Talking About?

In 2018, the National Academies of Sciences, Engineering and Medicine held a multistakeholder workshop to identify challenges and opportunities for the conduct of virtual clinical trials.[3] The workshop proceedings contain a tidy and unambiguous definition of what virtual trials are but seem to lack consensus on exactly what to call them. Virtual trials are defined as "…clinical trials in which all or part of the study incorporates digital health technologies and enables remote participation outside of the traditional brick-and-mortar study sites." Candidate umbrella terms for this kind of research were more heterogeneous, with "virtual" retained in the workshop title but "decentralized," "remote," "site agnostic," "direct-to-participant," "location flexible," "mobile," "flexible," and even "modern" and "21st century" suggested as possibilities by workshop participants.

The Clinical Trials Transformation Initiative, an organization with active

Figure 1. Traditional versus Virtual Research Approaches Contrasted.

Traditional Approach	Data Collected ...	Virtual Approach
Directly via observation/measurement	How?	Indirectly via connected devices
Brick & mortar study sites	Where?	Wherever patients roam or dwell
Patients & study personnel together	Who?	Patients alone (generally)
Research-specific data only	What?	Research-specific and/or 'personal' data
Prespecified intervals per protocol	When?	Prespecified intervals and/or continuously
For research purposes only	Why?	Sometimes for research, sometimes not

participation on the part of the US Food & Drug Administration, has released recommendations for what they refer to as "decentralized clinical trials," suggesting a preference for that terminology.[4] At this point, the terms "virtual" and "decentralized" are used more or less interchangeably, but as virtual approaches increasingly take root in the real-world setting, it is important to replace the term "trials" with "research" in recognition that the vast majority of real-world research is not trial-based. Hence, our use of the term "virtual research" throughout this paper.
To further establish exactly what we mean by virtual research, it is instructive to contrast it to traditional approaches in terms of a variety of questions related to data capture. This is summarized in Figure 1.

The how and where of data collection are fairly straightforward—in traditional research approaches, data are collected via direct assessment of study subjects at study sites, while virtual approaches eschew direct observation in favor of remote data capture via connected devices wherever patients happen to be. The who of data collection involves patients and study personnel together in traditional approaches, while patients are generally all alone in virtual studies (although there is some human interaction when telemedicine teams are utilized).

Finally, in traditional research the what, when, and why of data collection are all strictly guided by the study protocol, which governs that only research-specific data are to be captured, almost always at prespecified intervals. In contrast, things are more open in virtual approaches, as digital technologies capture research-specific data but also "personal" data along the way, and this can be done according to prespecified intervals or continuously. Indeed, in some instances, none of the virtually captured data were initially intended for research purposes, and this is important as we start thinking of classifying the different types of virtual research in the real-world setting.

A Classification Scheme for Virtual Research in the Real-World Setting

Real-world data sources can be distinguished along various dimensions, but for our purposes it is useful to focus on 2 in particular: one characterizing how the data are collected (active versus passive) and the other distinguishing the temporal aspect of data analysis (retrospective versus prospective).

Active data collection involves use of case-report forms, instruments or other means of data capture, where data are specifically collected for research purposes and patients are actively involved in sharing their data. In contrast, **passive data collection** refers to accrual of data in information technology systems as a by-product of real-world care processes or other patient activities. In this case, the data are not initially collected for research purposes but can subsequently be manipulated for use in research, and patients are not always mindful of the act of sharing their data.

The **prospective** versus **retrospective** distinction is straightforward, with prospective research involving the analysis of data collected from the present into the future and retrospective research involving analysis of data collected in the past.

Figure 2. Two-by-Two Typology of Real-World Data Sources, Highlighting Digital Technologies.

	Prospective Research	Retrospective Research
Actively Collected	Pragmatic Clinical Trials Noninterventional Studies Digital Technologies	Population Registries Digital Technologies
Passively Collected	Digital Technologies	Patient Charts (Hard-Copy) Databases (Claims & EMRs) Digital Technologies

When we combine these distinctions in a simple two-by-two typology (Figure 2), we can first see how the familiar real-world data sources (in black font) are sorted: pragmatic clinical trials and noninterventional studies such as registries in the upper-left quadrant; patient charts and computerized databases in the lower right; and population registries in the upper right. We also see that digital technologies (in red) appear in all 4 quadrants as a source of real-world data.

This enables us to start distinguishing different kinds of virtual research:

Actively Collected/Prospective Research. These include studies where connected devices are used to measure "novel endpoints" in both interventional and non-interventional prospective studies. In all other respects, these studies are similar to traditional prospective studies in that they require ethics approval, informed consent, a protocol to govern data collection, the whole nine yards.

An interesting example of this kind of real-world research is the "Cloudy with a Chance of Pain" study, which piloted an app designed to assess associations between weather and joint pain in patients with rheumatoid arthritis.[6] Participants entered self-reported pain, fatigue, physical activity and other data into the app on a daily basis for 60 days. Global positioning systems (GPS) embedded in their smartphones enable linkage to local weather conditions, thereby allowing weather data to be pulled into the study database and matched by time and location to patients' symptom data. Analyses of these data assessed associations between weather data and various measures of chronic pain, and found that higher relative humidity and wind speed and lower atmospheric pressure were associated with increased pain severity in people with long-term pain conditions.[6]

Actively Collected/Retrospective Research. These studies require de novo creation and curation of a database, are guided by a protocol, and require identification and recruitment of a study cohort and arrangements for data collection via digital technologies. Ethics approval and informed consent are required, as is the case with traditional population registries.

An interesting example of this type of real-world research is the "All of Us" population-based research program that is seeking to enroll a diverse group of at least 1 million people in the United States to accelerate biomedical research and improve health.[7] Elements of the protocol include health questionnaires, electronic health records, physical measures, and the collection and analysis of biospecimens. Although not an example of fully virtual research, study participants have the option to contribute data from their wearables and sensors. The program launched in May 2018; one year later, the program had met more than one-fifth of its recruitment goal.

Passively Collected/Retrospective Research. In this type of research, data flow automatically to the device/app developer without a protocol and with no active patient involvement. Consent for data sharing is handled via opt-in at the time of device/app registration. No formal ethics approval is required, nor is any advance work on the part of the researcher. The most common type of real-world data in this category derive from wearables, which have the capacity to continuously transmit data back to the study database without active engagement on the part of the wearer.

An example of this research is the Fitbit Sleep Study, which tapped Fitbit's longitudinal sleep dataset—built from millions of nights of data obtained via its Sleep Stages app—to determine how age, gender, and other factors affect sleep quality.[8] The Sleep Stages app uses motion detection and heart rate variability to estimate the amount of time users spend awake and in light, deep, and REM sleep each night. Data flow automatically to the database on a nightly basis, thereby leading to an ever-expanding dataset accessible for use by researchers, all of which occurs without any overt effort on the part of Fitbit users.

> The digital revolution in health is invading the clinical research realm, and nowhere is this invasion more pronounced than in the real-world setting. The COVID-19 pandemic has acted to accelerate these developments on all fronts.

Another, more timely example that has gained prominence during the COVID-19 pandemic derives from data collected by smart thermometers. One manufacturer of these thermometers, Kinsa, has created a website containing a heat map of elevated temperature readings derived from users of their device, which utilizes GPS technology to aggregate average temperature readings across the United States.[9] Historically, elevated temperature readings have been a leading indicator of flu outbreaks and now do the same for COVID-19.

Passively Collected/Prospective Research. In this type of study, data flow automatically to the device/app developer or to the study database (if separate), with no active patient involvement. In this instance, however, a protocol is required for identification and recruitment of the study cohort, and arrangements for data collection via the app(s) and device(s) involved. Ethics approval and informed consent are required.

The Apple Heart Study, a prospective observational cohort study, that has enrolled more than 400,000 participants to test the ability of a smartwatch algorithm to identify pulse irregularity and variability that might reflect previously undiagnosed atrial fibrillation.[10] Patient screening, consent, and data collection all happen electronically via an accompanying smartphone app, and the only thing that participants are required to do in the study is wear their Apple watches. Additional patient engagement and data collection are undertaken only for those participants in whom irregular heart rhythms are observed.

This simple classification scheme demonstrates how digital technologies fit in with other real-world data sources and facilitate greater understanding of different kinds of virtual research in the real-world setting. Some virtual studies will be more like traditional prospective observational research—and therefore take on the characteristics of registries, for example—while in other instances, real-world data collected by means of wearables and other connected devices will be tapped into for retrospective analyses, in much the same way claims databases have been for the past few decades. Recognizing these differences is essential to fully appreciating the nuances of virtual research in the real-world setting.

Challenges in Virtual Research Execution

In addition to presenting challenges to real-world research design, virtual approaches involve a host of challenges in study execution.

Not surprisingly, these challenges derive from the elimination of study sites and the critical role that site-based staff play in the research process. Here are 3 broad challenges that virtual approaches impose on study execution:

(1) **Patient recruitment**: How to identify potential study subjects without investigators to refer their patients and without site-based personnel acting as intermediaries and facilitators.

(2) **Ascertainment of eligibility**: If patients complete screening forms remotely, by themselves, how to ensure that they actually meet key eligibility criteria for study participation without corroboration from study sites.

(3) **Assurance of patient reliability**: How to get patients enrolled, stay engaged, and complete data collection without site support.

As these issues make clear, virtual research puts a far greater onus on patients to drive the success of the study—so we can see that patient centricity carries with it increased patient responsibility in the research process.

Fortunately, the same technologies that make virtual research possible provide solutions to the implementation challenges to which virtual approaches give rise. Patient identification can be facilitated by geo-targeted digital recruitment, such as pop-up ads on social media outlets and internet search engines. Patient eligibility can be ascertained by including electronic medical records access in the consenting process, thereby permitting the study team to contact the patient's healthcare provider to confirm diagnosis, medical history, medication use, and the like. And smartphone apps can be programmed with reminders and gamification elements to ensure that patients continue to transmit data and stay engaged throughout the study duration.

The Road Ahead

The digital revolution in health is invading the clinical research realm, and nowhere is this invasion more pronounced than in the real-world setting. The COVID-19 pandemic has acted to accelerate these developments on all fronts. Manufacturers remain cautious about deploying virtual approaches in their phase II-III clinical trials and have come to view the real-world setting as a lower-risk testing ground for innovation. Understanding how real-world data derived from connected devices compare to the sources we are already well familiar with is critical to sound study design—just as we readily discern a database analysis from a registry study, we should similarly be able to distinguish between different types of virtual research. For now, during this nascent phase of virtual research, the simple two-by-two typology described in this paper may prove useful, but look for it to give way to more complex classification schemes as further examples of virtual approaches proliferate in the real-world setting.

References

1. Center for Information and Study on Clinical Research Participation (CISCRP). Charts & Statistics. https://www.ciscrp.org/education-center/charts-statistics. Accessed May 1, 2020.

2. Transparency Life Sciences. XTalks Webinar: "Virtual Clinical Trials Are Here: Now What?" https://xtalks.com/webinars/virtual-clinical-trials-are-here-what-now/#description. Accessed May 1, 2020.

3. National Academies of Sciences, Engineering, and Medicine. 2019. *Virtual Clinical Trials: Challenges and Opportunities: Proceedings of a Workshop.* Washington, DC: The National Academies Press. https://doi.org/10.17226/25502

4. *Clinical Trials Transformation Initiative. Decentralized Clinical Trials.* https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/dct_recommendations_final.pdf. Published September 2018. Accessed May 1, 2020.

5. Reade S, Spencer K, Dixon WG. Cloudy with a chance of pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. *JMIR Mhealth Uhealth*. 2017.;5(3):e37. Published online 2017 Mar 24. https://doi.org/10.2196/mhealth.6496.

6. Dixon WG, Beukenhorst AL, Yimer BB, et al. How the weather affects the pain of citizen scientists using a smartphone app. *npj Digital Medicine*. 2019;2:105. https://doi.org/10.1038/s41746-019-0180-3.

7. The All of Us Research Program Investigators. The "All of Us" research program. *N Engl J Med* 2019;381:668-676.

8. Kosecki D. How Much Sleep Do Fitbit Users Really Get? A New Study Finds Out. FitBit. https://blog.fitbit.com/sleep-study. Accessed May 1, 2020.

9. US Weather Map. https://healthweather.us. Accessed May 1, 2020.

10. Turakhia MP, Desai M, Hedlin H, et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: The Apple Heart Study. *Am Heart J.* 2019;207:66-75.