



Published in final edited form as:

Sci Transl Med. 2015 July 22; 7(297): 297ed10. doi:10.1126/scitranslmed.aab1206.

App-enabled trial participation: Tectonic shift or tepid rumble?

Stephen H. Friend

President of Sage Bionetworks, Seattle, WA 98109, USA

Apple's Chief Operating Officer, Jeff Williams, surprised crowds at the spring launch when he revealed "ResearchKit," a collection of iPhone apps designed to allow individuals to collect their clinical data—and contribute to the precision medicine movement—outside the confines of hospitals and labs. But are these simply a smattering of souped-up health apps in a sea of thousands (that is, no big deal)? Are they support tools for uncontrolled clinical trials, which won't produce meaningful results (not to mention superfluous, given that patient-centered outcomes initiatives are well under way)? Or are they precursors heralding a tectonic shift in how people participate in their health management as well as in human disease research and clinical trials? The answer might depend more on human psychology than human health science.

CLINICAL TRIAL CONUNDRUM

Much of our understanding of the effects of modulators (such as drugs) on human diseases comes from clinical studies. Today, tens of billions of dollars are spent on clinical trials that range from large longitudinal observational studies to intensive testing of potential new drugs. Trials are typically coordinated through physicians at specific institutions and primarily capture data via infrequent surveys or face-to-face transactions. The participants are termed "subjects" because data move in one direction—collection from the test subject to analysis by those who lead the studies. The subjects provide written consent through forms that state that the subjects agree to donate their data to the institutions, which then own the data. Thus human data obtained in such trials tend to be (i) held locally, (ii) difficult for scientists outside of the sponsoring institution to obtain and analyze, and (iii) derived from inadequate sample sizes—which all too often yields infrequently collected, noisy data. In an effort to address these issues, Apple recently worked with several institutions to develop apps that could allow virtually no-cost, fully scalable, sensitive, frequent collections of both data and insights from participant-centered trials in which the participants are expected to be the owners of their own medical data.

FIRST APPS

The first class of five apps relies on a software framework called ResearchKit that Apple has deposited as open source code on "GitHub" (1). Researchers at Mt. Sinai School of Medicine developed an app-based clinical study to examine individuals with asthma, to track both their symptoms and modulators of their disease. Scientists at Massachusetts General

Competing interests: The author was a part-time employee of Apple in 2015

Hospital designed an app-based clinical study for diabetes, while Stanford Medical School developed one that investigates ways to incent participants by allowing them to track aspects of their activities that might effect cardiovascular health. Sage Bionetworks collaborated with the University of Rochester to design an app-based clinical study for Parkinson disease, and collaborated with researchers at the University of California, Los Angeles; Harvard University; and the University of Pennsylvania to initiate participant-centered trials with women after treatment for breast cancer to follow the resolution of side effects from radiation, chemotherapy, and hormone ablation.

After the first two months of use, several observations are worth noting. The use of iPhone-based apps has enabled broad enrollments (across the five apps, more than 70,000 participants enrolled) throughout the United States (the asthma study now tracks state-by-state differences in the primary triggers for asthma, from anger to pollen) (2). Diverse variations are being seen in the effects of L-dopa among Parkinson patients when measured on a day-by-day basis (more than 10,000 comments have been collected). When collecting multiple dimensions of data from a given patient, the ability to build personalized classifiers that can track the effects of modulators and changes in the state of the disease become rather obvious. For example, it is possible to define a core set of 20 attributes most commonly found associated with Parkinson disease, and each patient displays a distinct cluster of these attributes. From these data, one can organize the most prominent attributes for each patient, and this cluster can then serve as a personalized classifier to be used both by the participant and researchers.

These app-based clinical trials should not be seen as pioneering the enrollment of participant-centered trials, as many such precursors exist—such as those supported by the government agency Patient-Centered Outcomes Research Institute (PCORI), which has supported several trials primarily focused on comparing evidence-based health care options (3). What is notable, however, is the rapidity of enrollment, largely by patients who indicate a desire to share their data broadly (75%). Also exciting are the potential benefits of shifting beyond surveys to the collection of sensor data that can be collected both directly and frequently with little hassle to provide streams of objective phenotypic data.

NEXT GEN APPS: ENGAGEMENT AND CONSEQUENCES

Despite some promising aspects already apparent from these ResearchKit-driven clinical app studies, there are several obvious issues to be solved. Most people who download apps use them for only short stretches of time before they get bored. Similarly, classical trials have well-known problematic issues related to patient retention (4). ResearchKit apps will need to invent ways to offer participants a creative experience that they can integrate into their daily lives and from which they both benefit personally and gain a sense of being a part of something larger than themselves that has the potential to help others. Participant forums—such as chatrooms with researchers—or immediate feedback that gives, for example, a running tally of participants in one's hometown or demographic or a link to a related new scientific paper or personal story might provide the needed push to participate.

There are strong selection biases among those who participate through the use of a defined product, such as an iPhone, that will affect the ability to reflect these findings across broad segments of populations. Furthermore, participant-centered trials have, as a core feature, the sharing of data by participants about themselves. Although there is great value in people contributing data about themselves, we anticipate that such knowledge can bias participants, causing a shift in the calculated effects of health modulators; this assumption is in keeping with the placebo effect and is a reason why double-blinded trials are popular. Psychologists have long been aware of the significant potential contamination when participants and researchers are not blinded to the study components (5).

NEXT DIRECTIONS

The use of pervasive computing devices, such as smartphones, as the vehicle with which to collect data and insights has as one of its main advantages the ability to shift away from simply collecting subjective data combined with infrequent collections of objective data. Instead, biosensors allow one to collect real-time objective data about a disease that would normally be collected within subjective assessments—such as “How would you rate your recent pain level?” (which is prone to vague recollections). As the “Internet of things” (6) (including smartphones) emerges, we can realize the power of real-time data. Before that can occur, we need to learn how to parse signal from noise and to validate these new measures.

Existing markers of disease, such as cholesterol levels and blood pressure, were developed by linking the numerical levels to disease progression. When a new candidate biomarker is developed, it also must be linked to disease progression. As with any new biomarker, the existing validated health measures that we are now capable of following with biosensors rarely have preexisting biomarker data that could be exploited to perform the needed comparisons. This means that, for each symptom that could help drive a ResearchKit study, we will need to wait for an interval of time to pass during which we can collect both the well-validated existing measure and accompanying new sensor data to complete the so-called “validation loop.” A validation loop refers to the process of identifying a new candidate biomarker and then linking changes in the biomarker with changes in the designated disease.

Furthermore, there is the exciting but puzzling possibility that real-time measures from sensors might highlight day-to-day variations that current measures had assumed were simply noise. For example, if every text or e-mail we type can be used to track cognition fluctuations, then we might learn how we need to change our experimental designs to truly study cognition rather than using existing cognitive tests that pride themselves on their lack of being influenced by diurnal variations.

One approach being used in some of the ResearchKit studies is to build three-layered stacks of information that tie together (i) reasonably well validated survey questions that can only be administered infrequently with (ii) structured tasks that have a defined activity for a short-defined period of time and (iii) passively acquired continuous feeds of data. As an example, in the mPower app to follow Parkinson disease patients, there are infrequent surveys for tremors and movements, structured tasks such as for tapping and walking, and continuous

location data feeds that provide position without divulging exact locations. This three-layered stack of data is being used to help ground the passive and the structured-task data so that, combined, there is a better chance of making sense out of the eventually powerful feeds of continuous passive data. This approach of linking high-friction, well-validated surveys to moderate-friction structured tasks and the inclusion of passive data streams could be used to help track a wide range of symptoms. As this practice becomes more common, there could be a powerful transition of classifying diseases by isolated symptoms as might be collected in a physician's office to the real-time collection of variations in daily activities that better reflect the full dimensions of diseases when revealed by the multiple dimensions of real-time data collected from sophisticated sensors. This advance will require the mapping of current symptoms and their paired variations in daily tasks back onto the genomic defects associated with various diseases. Such ideas will require large cohorts, some of them very well phenotyped and genotyped, and a way to navigate through the impending morass of related apps that all will vie for becoming the standards by which to follow the symptoms underlie various diseases.

The current ecosystem of academic scientists and start-up companies each hoping to build out ways to follow various symptoms is unlikely to be an efficient way to transition to a set of standard ways to follow the diverse symptoms that define health and disease. This then begs the question—what will it take to accelerate the uncovering of robust ways to track various symptoms and adoption of them as standard tools?

Currently, ResearchKit code (1) and codes for the first class of apps have been made available as open-source code on GitHub. Mechanisms for individuals to work as a group to build new symptom modules (such as for cognition or mood) for ResearchKit and for enticing these groups to make widely available their raw data, code, and analyses could accelerate the adoption of a standard way to assess that symptom. There exists a long history of nurturing such collaborations, as was done by astronomers (7) and by the SNP consortium for DNA variations (http://metadatabase.org/wiki/the_SNP_consortium_database). Such “federated” approaches to developing new symptom modules could be very efficient but would need to somehow be nurtured by interested parties, possibly including funders and participants.

ULTIMATE ECOSYSTEM

As we anticipate a world in which data and insights surrounding aspects of our health and disease become more available through the use of pervasive computers and the “Internet of things,” it is likely that there can be one further acceleration in the existing biomedical paradigm. Efforts in translational medicine have primarily been driven by a linear process of designing a study to ask a pertinent question about health and disease, finding someone who will fund one's efforts to generate the ideas, analyzing the data, and publishing the findings so another person can take a turn on the crank. Long delays are possible at each stage, but the dominant one is that we assume that most data are generated for the question being asked. When this kind of delay applies to a longitudinal study, the turn of the crank could be measured in years to decades. The use of real-time sensor data that can be mapped onto symptoms collected among individuals and that go beyond being condensed into standard,

well-defined signs and symptoms means that one might be able to analyze biomedical data as if each of our lives were continuous longitudinal studies. If the accumulation of data were consented by app-driven participant-centered trials so that the resulting body of data was available to qualified researchers worldwide, then the time it takes to iterate and discuss new ideas might become equivalent to the time required to perform analysis of the data.

Recognizing the wisdom of Yogi Berra's saying, "It's tough to make predictions, especially about the future," I'll bet that real-time streaming of data from pervasive computing devices, as has been shown possible through the participant-driven trials enabled by the ResearchKit apps, may be equally poised to impact precision medicine as the efforts to make legacy medical records interoperable.

REFERENCES AND NOTES

1. ResearchKit code. available at <https://github.com/researchkit/researchkit>
2. Asthma health study. Icahn School of Medicine; Mt. Sinai: available at <https://itunes.apple.com/WebObjects/MZStore.woa/wa/viewSoftware?id=972625668&mt=8&ls=1&v0=www-us-researchkit-itms-asthma-health>
3. PCORI. www.pcori.org
4. Cooley ME, Sarna L, Brown JK, Williams RD, Chernecky C, Padilla G, Danao LL. Challenges of recruitment and retention in multisite clinical research. *Cancer Nurs.* 2003; 26:376–384. quiz 385–386. [PubMed: 14710799]
5. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: Recent advances and current thought. *Annu Rev Psychol.* 2008; 59:565–590. [PubMed: 17550344]
6. Atzori L, Iera A, Morabito G. The internet of things: A survey. *Comput Netw.* 2010; 54:2787–2805.
7. Seamless astronomy. available at <http://projects.iq.harvard.edu/seamlessastronomy>