LANDSCAPE:
Overview of Digital Technologies

In 2013, 91% of US adults owned a cell phone. The penetration rates are high across ethnic and racial groups, as well as across ages. In early 2013, the Pew Foundation’s Tracking for Health study found that 69% of Americans monitor some form of health related information, and 21% of them use some form of digital device to do so (Fox & Duggan, 2013). This high rate of adoption for digital technologies presents profound opportunities to more intensively monitor health outcomes, better characterize the behavioral and environmental influences on health and illness, and to intervene frequently and in context to improve health. Even a basic cell phone (a “feature phone” rather than a smartphone) can be used to engage participants, assess outcomes, and deliver messages. They can be used to give and receive data via voice and SMS text messaging, identify location, and provide estimates of the level of social contact and connectedness. These opportunities expand exponentially when using a smartphone, tablet or e-reader, which has substantially more computing power, applications, interface capabilities, and internet access than feature phones. Smart-phones are owned by 60%+ of US adults and are the fastest growing segment of wireless communication devices. (Smith, 2013)
**Smartphones:** Smartphones are a natural point of engagement for much of the US population. Data suggest that for adolescents and young adults, digital data connections are now more frequent than voice calls (Pew Research Center, 2015). Even for older adults, the rate of smartphone adoption is growing rapidly (Smith, 2014). The phone can be used for voice, SMS text and email, but also through pre-recorded video and video chats (e.g., Skype). This can facilitate a wide range of remote, interactive communications with participants, including remote consent using technology to reduce the variability and enhance the understandability of a complex process (Kumar et al., 2013a).

Smartphone internal sensors can also collect data by capturing physical movement (from GPS, gyroscopes and accelerometers), location and transportation mode (GPS), sound, images (camera), social interactions (through location, text and voice) and some physiological parameters (Kumar, Nilsen, Srivastava & Pavel, 2013b). Smartphones are also capable of assessing physiological parameters when the participant interacts with the phone (e.g., pointing the camera at their face, pressing their finger on the screen, blow into the microphone, etc.; Kumar et al., 2013b). These data can be used to infer the level and type of physical activity, sedentary behaviors, sleep, and mood. Physiological data, such as heart rate can be collected via the phone’s camera or through sensors that can be worn and that interface with the smartphone. Patient Reported Outcomes (PROs, e.g., PROMIS, Cella et al., 2009) and performance measures (e.g., NIH Neuroscience Toolkit, finger tapping, memory tests; e.g., Zelazo et al., 2014) can be easily and frequently deployed through apps, emails, texts, voice actions, and experience sampling or Ecological Momentary Assessment (EMA) methodologies that can provide these assessments via fixed or random schedules or in response to events (e.g., query when in a specified location, when a physiological threshold is exceeded). Finally, data collection and aggregation can be done through apps. These apps can be custom-made for projects or use data collected from the APIs of existing apps.
Many of the capabilities of smartphones are being integrated into smartwatches. These smartwatches have the advantage of being in touch with the user’s skin continuously for potential continuous physiological assessments such as heart rate, oxygen saturation, temperature, skin conductance (Kumar et al., 2013b). Smartwatches also provide wrist accelerometry which can be used to more precisely identify a range of activities (e.g., smoking, eating, walking, running) as well as estimate energy expenditure (Kumar et al., 2013b). In addition to phone-based sensors and smartwatches, other wearable devices and sensors that connect through smartphones or directly to the Internet via wireless communication are also emerging. These are often used with stationary sensors (such as blood pressure cuffs and weight scales) with the data transmitted through Bluetooth on the phone or via home-based wireless networks to data stores in the cloud or repositories of interest. Data can also be collected through devices such as the Microsoft Kinect which uses infrared sensing to track movement (e.g., Galna et al., 2014) and can be used to validate mobile data in the home.

The wearable device ecosystem has had an explosive growth (470+ wearable health trackers); most are tracking activity, sleep, and heart rate. This includes a growing range of commercial sensors that vary in quality, cost and focus, but which are now owned by
an increasingly large proportion of the population (Fox & Duggan, 2013). Sensors also
have been developed by industry specifically to target medical conditions. Products such
as AliveCor (ECG; Lowres et al., 2014) and BioRythym (Activity, BP, HR, hydration and
metabolism) all offer a range of options for measuring specific biomedical indices
throughout the day. Sensors that are FDA approved must show evidence of validity and
safety. Many of these wireless sensors are designed for specific conditions, such as
diabetes and asthma and are already deployed in specific patient populations (e.g.,
glucometers, spirometers). In addition to commercial sensors, there are also a number
of finely calibrated sensors designed for research. These sensors target exposures like
sun exposure (Buller, Berwick, Lantz, Klein Buller, Shane, Kane, & Liu, 2015) or second
hand smoke (Liu, Antwi-Boampong, BelBruno, Crane, & Tanski, 2013). Environmental
exposures play a major role in health and may trigger epigenetic changes or an
underlying genetic predisposition to disease (Jirtle & Skinner, 2007). In recognition of
this, in 2006 NIH funded the Genes, Environment and Health Initiative
(http://www.genome.gov/19518663). A major part of this program was funded research
to advance environmental monitoring, as well as person-level exposures to
environmental contaminants. The results of this research led to a family of new tools
and technologies that improve the ability to measure person-environment exposures
(e.g., activity, dietary intake, smoking, particulate matter exposure)
(http://www.genome.gov/19518663). Finally, there is a growing field of medication
adherence sensors including glow caps, capacitive sensors, and Proteus, an ingestible
sensor embedded in pills that verifies their consumption (Kane, Perlis, DiCarlo, Au-
Yeung, Duong, & Petrides, 2013). All of these can be deployed in ways that transmit data
in real time via the Internet (WiFi) or Bluetooth.

Digital trace data: Complementing data that is driven by personal mobile devices, data
about individuals are being captured passively as people communicate with one another
on social networks, shop, work, or engage in any activity that is capable of leaving
have been published on everything from obesity (Christakis and Fowler, 2007) to
psychological distress during the recent economic recession (Ayers, 2012). Private
industry has capitalized on this trend to refine and personalize services and marketing,
often to a remarkable degree (Turow, 2011). These data support understanding of
human dynamics and social interactions in near real-time. Mining these data for insights
about the role of social connections and influences on health is an enormous
opportunity. This opportunity is much more than the current text mining of Facebook
and Twitter posts, and includes an increasing interest in mining image and other types
of data for trends in affect, social interactions and environmental influences. These
digital footprints are ubiquitous and can be obtained unobtrusively without respondent
burden, either anonymously or with participant permission.

Scientists in academia and industry are engaged in highly promising research aimed at
miniaturization and fusion of wearable sensors that, in turn, will have the capability to
provide parallel, simultaneous streams of continuous data on multiple health-relevant
parameters (Davies, 2013). All of these data will allow for a unique opportunity to capture intensive longitudinal data that can characterize not only risks and exposures, but also changes in health status, occurrence of behaviors of interest (e.g., stress, smoking and other addictive behaviors, medication usage, diet, physical activity, fatigue, sleep, and social interactions), and individual responses to exposure. Leveraging genetic data from the cohort, these data can be used to refine behavioral phenotypes, evaluate potential endophenotypes, and assess gene-environment interactions.

These technologies also can be leveraged to capture the effectiveness of interventions in what would be much closer to a ‘real world effectiveness context’ than typical efficacy trials. Further, strategically deploying evidence-based technological adherence interventions may support treatment studies by enhancing medication adherence and identifying adherence issues before they disrupt a trial. The data from mobile and wireless technologies also provides a unique opportunity to capture individual’s responses to various treatments in real time. This temporally-dense health outcome data can potentially reduce the number of people needed in a clinical trial, reduce the length of trials, identify and reduce adverse events by detecting them sooner, and enhance participation through reminders and other engagement strategies (Kumar et al., 2013a). Technology is critical for participant engagement because consumer electronics can be leveraged to connect and partner with volunteers, on a schedule they control, and enhance participation by providing access to personal, trusted health information in a secure, user-friendly way. Further, the way in which users interact with the system provides data that can used to iteratively refine and enhance the usability and usefulness of the system. It can also provide an affordable and reliable communication system between the study team and the volunteer.

**WORKING GROUP RECOMMENDATIONS:**

1. **Scope of mHealth for the cohort** – By the time a national cohort can be launched, the penetration rate of smartphones should be sufficient to use smartphones as the primary portal for data collection and participant engagement. Use of the smartphone will allow for use of a variety of internal sensors in the phone, messaging capabilities of the phones (e.g. SMS) and a range of special purpose sensors that can interface with the phone and transmit health data wirelessly. Depending on rates of use by the initiation of a national cohort, smartwatches also should be considered, at least for a subset of the cohort given the additional physiological and activity data available via smartwatches. Given the extensive range of possible variables collected by these technologies, it is important that what is collected be driven by the scientific questions posed, not by what is possible to collect from these devices. mHealth technology can be utilized for:

1. Cohort enrollment, consent, engagement, and to facilitate retention
   - Tools for consent are now available or in development, e.g., Sage’s e-consent
   - Software that can return information to participants, providing immediate feedback or requests for input.
II. Technology-based longitudinal data collection to better characterize medical outcomes, including:
   • Patient-reported outcomes of symptoms and functioning,
   • Performance tests (sensory, motor, cognitive),
   • Measures for specific conditions (e.g., weight scale, glucometer, spirometer, blood pressure, which now often have a wireless interface to avoid repeated manual entry.
   • Sleep monitoring

III. Technology-based longitudinal data collection currently available to better characterize treatments, including medication adherence monitoring.

IV. Technology-based data collection to characterize behavioral and environmental exposures (probably in a subset of entire cohort given costs and effort required)
   • Personal particulate matter exposure (with other particulate matter exposure obtained from a combination of cell phone locations, durations, and EPA data for areas)
   • Physical activity, sedentary behavior, duration of moderate and vigorous physical activity, characterization of daily activities
   • Diet monitoring (prompted food diary recordings, 24 hr. recall, or pre-post meal pictures)
   • Smoking (smoking sensors on those who indicate smoking)
   • Secondhand smoke (home sensors on those who indicate living with those who smoke)
   • Stress responses (heart rate variability, galvanic skin response-GSR)
   • Social behavior (time alone, number of social contacts per day, etc.)
   • Geolocation to estimate a variety of environmental exposures and social determinants of health

V. Note that any of these measurements can be performed on a regular schedule, a randomly prompted schedule, or prompted in response to a specific event (e.g., when at a specific location).

2. Consumer vs. Research Grade mHealth - There are thousands of health apps, but these are generally ignored by the research community because most are not evidence-based or experimentally validated. A prime example of this is Abroms and associates (2013) exploration of the quality of commercial apps for smoking cessation, which found almost no evidence of evaluation or even use of best practices. This is in contrast to Free and colleagues (2011) work to generate an empirically-supported mobile smoking cessation intervention. The same is true for research vs. consumer grade sensors. Actigraph, for instance, has solid validity data for measuring physical activity (Evenson & Wen, 2015), while most commercial accelerometers (e.g., FitBit) still need validation for research purposes (see preliminary data in Case, Burwick, Volpp & Patel, 2015). For some devices, this is difficult to do because of the proprietary nature of the data and algorithms. This does not necessarily mean that commercial sensor data are unusable, but validation of these devices against the research grade standards is necessary before
many of the commercial sensors can be used in research. Many of the commercially available sensors for specific medical conditions, e.g., glucose monitors for diabetes, have already been tested and approved by FDA. There are benchmarks in place for new sensors in these areas.

3. **Assessment vs. Intervention** – Cohort studies often experience a tension between observational versus clinical trials research. Given the capabilities of mobile technologies, this tension is particularly salient for mHealth research. Scientifically, the methods for validating a sensor are no different than validating a biomarker and many research grade sensors have been validated, at least in laboratory settings. Emerging sensor technologies not yet validated could be rapidly validated within the context of a large cohort study.

In contrast, interventions delivered on a mobile platform have more limited clinical outcome data to date, but that is in part due to the long time frame for conducting a randomized clinical trial (RCT). If an investigator submitted an R01 grant to develop and evaluate a smartphone app as soon as the iPhone was released (2007), the results of this research would likely be published only recently (Riley, Glasgow, Etheredge & Abernethy, 2013). While the cohort could be used as a test bed to evaluate certain mHealth interventions, the primary use of mHealth in this cohort would be for better assessing disease outcomes and for assessing behavioral and environmental exposures.

4. **FDA Regulation**: Unless used for diagnostic or medical treatment purposes, FDA regulations do not apply to mHealth sensors and interventions. Further, the FDA has clear guidance on regulations for mHealth that were updated in February 2015 (http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf). The FDA considers many of the mobile apps and devices that would be utilized for a cohort study as either exempt or apps for which it has chosen to exercise discretion. For those devices or apps that require FDA approval, there are both assessment (http://www.alivecor.com/press/press-releases/alivecor-receives-first-fda-clearance-to-detect-a-serious-heart-condition-in-an-ecg-on-a-mobile-device) and intervention (http://www.businesswire.com/news/home/20130613005377/en/WellDoc-Launches-BlueStar-FDA-Cleared-Mobile-Prescription-Therapy#.VN1PdkJ0GHw) apps, sensors and mobile systems that have been approved. IRBs are becoming increasingly well-versed on mHealth assessment and intervention protocols, and there is a considerable literature on privacy and security issues (Martínez-Pérez, de la Torre-Díez, & López-Coronado, 2014).

5. **Validation Cohort**: As noted in other areas, cohorts can be used to deploy and validate emerging technologies and to cross-calibrate existing measures with new measures to maintain consistent data outputs over time while remaining current with the latest technologies. Existing cohorts already collect data that is considered the gold standard, and the new technology can be compared to these benchmarks without the full cost of new trials.
6. **Member Engagement:** This is a critical element of a mobile-enabled cohort. Mobile allows for user-friendly interactions for both receipt and transmission of data. Thus, creating a Precision Medicine Initiative member engagement app that would provide participants with data control, prompts for data entry, immediate feedback after data entry, etc., is a feasible and manageable project. The design of the app should support a low burden for the participant and passively return information and indicators (as appropriate) to participants. User-interface and technology design research can be brought to bear on this issue (e.g., http://www.usabilitynet.org/management/b_overview.htm). There is a science for developing technological tools for diverse users (e.g., the disciplines of Human Computer Interaction and User-Centered Design and User Experience). The work done for the cohort in this area can help set the standards for mobile monitoring in other cohort studies. Given the costs of traditional participant engagement, this may be one of the best uses of cellphones for this project.

**Barriers/Challenges**

- **Standards for Sensor Data:** There is currently no standardized data format to share de-identified information; most vendors have a closed ecosystem locking in consumer data. There are efforts underway to standardize mobile data (e.g., Continua, HL7)

- **Privacy, Security and Verification:** There is growing concern that the wide range of health data collected using cell phones may not be secure (Urban, Hoofnagle, & Li, 2012). These security issues are often user-generated data leaks (e.g., lost phones without passwords), but malicious loss may increase as the phone becomes the home of financial information. Data collected must also be easily and reliably identity-verified and not intentionally or accidentally collected if devices are used by other individuals. Privacy and security is a focus of an upcoming IOM workshop, a multi-federal group led by the White House and the Secure and Trustworthy Cyberspace program at the National Science Foundation (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504709). Findings from this research can serve as a basis for insuring appropriate security and privacy in a mobile-enabled cohort.

- **Motivation:** A well-designed incentive system will be needed to maintain consistent study participation. There is a wide range of possible incentives, ranging from financial to social (e.g., Facebook ‘likes’ for organ donation) to provision of information that is perceived as useful (e.g., relevant personal health information). This will have significant overlap with the volunteer engagement section, but is an extremely important aspect of the cohort effort that can be
facilitated by mobile technologies that can deliver these incentives immediately following a participation event.

- **Data Quality**: The quality of the data must be understood and documented, which is critical for interpretability. Many factors affect data quality, including the original rationale for development (medical versus entertainment, etc.), the way in which it is used, or other factors associated with the user and the operating context. Deploying the devices in a way that ensures high data quality is essential. Initially, this means limiting the devices initially accessed to ensure the value of the data being collected are known. These capabilities can expand over time.

- **Ease of Use**: Technology for wide deployment must be used by people with varied technology backgrounds, socioeconomic and educational factors, and medical conditions. Addressing this issue will require working with the use community and learnings from the human-computer interaction and user-centered design disciplines.

- **Connectivity**: Sufficient connectivity and bandwidth must be largely available for broad adoption (currently problematic in rural areas). This area is currently a focus on the FCC’s rural broadband project (http://www.fcc.gov/encyclopedia/rural-broadband-experiments)

- **Exposure assessment**: The state of the art of measuring exposures to environmental insults is still nascent. Needed are better mobile devices that can be worn continuously to capture data on air pollution, noise, allergens and other toxins (indoors and outside) that have known or hypothesized influences human health. This area has been a focus of the National Institute of Environmental Health Sciences at NIH (http://www.niehs.nih.gov/)

- **Representativeness**: Representativeness of the data is a potential problem if participation is limited only to those with smartphones. How do we ensure that individuals are enrolled in cohorts in ways that allow generalization of at least some of the knowledge gained? This is an important challenge vis a vis high-risk patients with multiple chronic diseases who may not be early adopters of technology, although recent research suggests that an increasing number of people with serious illness are adopting technology, in part to manage disease and treatment information. This information, paired with the rapid rate of adoption of smartphone technology in the population, suggests that this issue will become less important over time.

- **Smartphone compatibility**: Need to ensure compatibility of data collection apps with the myriad devices that are diverse in hardware, operating system type, and
even version of operating systems. Upgrades in operating systems and hardware also need to be anticipated, tested, and made available to the user when these upgrades occur.

- **Rapid proliferation of new technologies:** Many of the technologies currently available for this project did not exist a decade ago, and currently unimaginable technologies are likely to be developed in the next decade. These advances will need to be incorporated as the technology develops and co-calibration or backward compatibility of data from earlier to current technologies needs to be insured.

- **Cost of devices:** Cost of sensors can grow very quickly and become prohibitive, especially as wireless wearable biosensors to capture exposures that cannot be assessed using stand alone smartphones are added.

- **Ownership of the data:** Many forms of “personal health data” (Clarke et al., 2007) derived from traditional and digital trace data sources are owned by commercial entities that are not primarily interested in public health. Many of these entities have a proprietary interest in not sharing information on the validity or efficacy of their products. Obtaining access to data from commercial entities on behalf of their users would be much more efficient than going to each end user individually, but obtaining data access from these entities may be problematic. To be maximally useful, these data would need to contain identifiable information to allow linkage with other data sources which makes the consent process more complicated.

**AUDACIOUS IDEAS**

Opportunity to link mobile health data with genetic data, health care utilization, real-time environmental exposure data, and various social and behavioral determinants of health to develop better health outcome prediction models.

- Recruit sufficient numbers of volunteers who are willing to open up essentially all of their data – genome, microbiome, medical/EMR, behavioral, social, precise geospatial (via addresses & GPS data) - so that the multiple influences between and among these can be understood. In some cases this would involve entire households, families, and communities making their data available to understand how these relationships impact the life-course of individual health. This could be based upon a movement to “donate your data” for purposes of science.

- Create a new public culture of motivated participation and partnering in one’s healthcare care for an individual’s own benefit and for research that will improve the overall health of the nation/healthcare enterprise.
• Accomplishing what is described in the first bullet could be facilitated by the “Synch for Science” (S4S) initiative described in the Precision Medicine Electronic Health Record workgroup background paper. The EHR group makes a strong case for the difficulties in surmounting the many challenges of separate medical record systems, IRB boards, institutional norms for providing patient data and the like. They suggest that the already authorized Blue Button functionality could bypass this with a campaign that would encourage patients to download their data and then voluntarily provide it to a repository that would support Precision Medicine research.

• Promote competitions between communities across the US to become “Framingham 2.0”. This could involve a social movement that engages multiple stakeholders at every level – health, education, business, faith-based, philanthropic and others – so that a culture of data sharing is promoted.

• “Simple” risk factor models can be improved substantially to reflect genetic, molecular, behavioral, social, and environmental determinants of health in ways that facilitate more effective interventions and potentially improves analysis of interrelationship between co-morbidities.

• Sophisticated exception handling process for data allows health systems to be paid based on population health risk, shifting the focus to preventing disease and disease exacerbations among high risk patients as opposed to the current patient-initiated care seeking paradigm.

• In addition to preventing events through better addressing risk factors, wearable devices could lead to preemptive medicine approaches such as fall prevention or early detection of imminent risk for myocardial infarction or a diabetic crisis.

• Epigenetics- study of the interplay between the environment (e.g. pollution, stress, chemical alterations in body etc.) and its effect on which genes are turned on or off (gene expression). Mobile health can play a significant role in understanding epigenetics, especially in the intensive longitudinal characterization of these environmental influences.
References


Buller DB, Berwick M, Lantz K, Klein Buller M, Shane J, Kane I, & Liu X. Smartphone Mobile Application Delivering Personalized, Real-Time Sun Protection Advice: A Randomized Clinical Trial. JAMA Dermatology, Published online January 28, 2015.


Smith A. Smartphone ownership-2013 update. Available at:


Urban JM, Hoofnagle CJ, & Li, S. (2012) Mobile Phones and Privacy. Available at:

Willett, WC Science 26 April 2002: Vol. 296 no. 5568 pp.695-698 DOI:
10.1126/science.1071055