The Pace of Technologic Change
Implications for Digital Health Behavior Intervention Research

Kevin Patrick, MD, MS,1 Eric B. Hekler, PhD,2 Deborah Estrin, PhD,3 David C. Mohr, PhD,4 Heleen Riper, PhD,5 David Crane, MSc,6 Job Godino, PhD,1 William T. Riley, PhD7

This paper addresses the rapid pace of change in the technologies that support digital interventions; the complexity of the health problems they aim to address; and the adaptation of scientific methods to accommodate the volume, velocity, and variety of data and interventions possible from these technologies. Information, communication, and computing technologies are now part of every societal domain and support essentially every facet of human activity. Ubiquitous computing, a vision articulated fewer than 30 years ago, has now arrived. Simultaneously, there is a global crisis in health through the combination of lifestyle and age-related chronic disease and multiple comorbidities. Computationally intensive health behavior interventions may be one of the most powerful methods to reduce the consequences of this crisis, but new methods are needed for health research and practice, and evidence is needed to support their widespread use.

The challenges are many, including a reluctance to abandon timeworn theories and models of health behavior—and health interventions more broadly—that emerged in an era of self-reported data; medical models of prevention, diagnosis, and treatment; and scientific methods grounded in sparse and expensive data. There are also many challenges inherent in demonstrating that newer approaches are, indeed, effective. Potential solutions may be found in leveraging methods of research that have been shown to be successful in other domains, particularly engineering. A more “agile science” may be needed that streamlines the methods through which elements of health interventions are shown to work or not, and to more rapidly deploy and iteratively improve those that do. There is much to do to advance the issues discussed in this paper, and the papers in this theme issue. It remains an open question whether interventions based in these new models and methods are, in fact, equally if not more efficacious as what is available currently. Economic analyses of these new approaches are needed because assumptions of net worth compared to other approaches are just that, assumptions. Human-centered design research is needed to ensure that users ultimately benefit. Finally, a translational research agenda will be needed, as the status quo will likely be resistant to change.


The Technologic Infrastructure for Health Is Changing

Digital technologies are increasingly pervasive in all aspects of daily life and the ease with which digital tools are adopted is in part because of the malleability and adaptability of digital technologies. The Internet and Web Architectures that underlie this digital revolution are fundamentally minimalistic and modular and allow for decentralized growth based on minimal commonality. Mobile and wireless technologies, including cellular systems, embedded cameras, smartphones, and the Internet of Things bring connectivity to every moment of one’s life. Search and online services that live on the Web continuously and automatically capture,
process, and adapt to analytics result in exponential and ongoing improvement in performance, functionality, and adoption. Technologic interventions in many areas, such as marketing or search strategies, are developed iteratively in ways that allow determination to be made of which version results in greater amounts of a desired user behavior. The promise of machine learning encourages recentralization, or at least federation, of data into highly scalable cloud infrastructures to support discovery.

For health and health care, the possibility of profound positive disruption is clearly present but faces challenges. Healthcare providers and technology giants are developing digital health products, services, apps, and platforms. The demand for apps is growing rapidly with approximately 165,000 self-labeled healthcare apps available in early 2015, but also the digital devices used routinely that provide a wealth of data from physiology to environment. These new data sources involve not only the technologies that individuals purposely use to monitor their health (e.g., smartwatches, commercial accelerometers, smartphone sensors and apps, heart rate monitors) but also the digital technologies used routinely that provide a wealth of data about behaviors, their influences, and consequences.

The digital traces left behind as people interact with their cell phones, social media sites, search engines, financial transaction systems, and everyday household items provide multilevel, temporally dense data on individuals and populations that can be utilized to advance understanding of human behavior. Soon, the Internet of Things will gather correspondingly dense physiologic data; for example, lavatories may soon be able to automatically perform biomarker and microbiota analysis and bathroom mirrors could be equipped with facial recognition software to identify health problems and breath sensors to monitor alcohol or tobacco use.

The process of discovery in this digital world is increasingly bidirectional with the advent of citizen science, described as “general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their tools and resources.” By assisting researchers in the analysis of big data and participating in large-scale experiments, citizen scientists are co-creating a new culture in which democratized research leads to greater and more rapid discovery. For emerging technologies that have very few users, services such as Amazon’s Mechanical Turk (www.mturk.com) enable recruitment of a diverse group of testers or participants for minimal cost. The rate of evidence and discovery is poised for tremendous growth thanks to new platforms such as Apple’s ResearchKit (www.apple.com/researchkit/).

Consumer-facing applications of these technologies bring many online and fuel improvements in the user experience and engagement. Although some studies suggest that things are improving, many digital health interventions experience high levels of attrition and as another paper in this series notes, securing the engagement necessary for lasting behavior change presents significant challenges for intervention developers.

These Changes Support the Ability to Handle Highly Complex, Multilayered Issues in Health

Concurrent with advances in information technology are three major trends in public health and medicine:

1. the emergence of chronic diseases as the main causes of poor health, disability, and death;
2. an increased understanding of the multiple influences on health, including the genome, microbiome, health behaviors, social influences, and the environment; and
3. collaborative, self, and social health management.

The combination of these poses both an unprecedented challenge to traditional health care and new opportunities for population health improvement.

The rising burden of chronic disease is a problem of both volume and complexity. Tobacco use, excessive alcohol consumption, diets high in sodium and low in fruits and vegetables, physical inactivity, and uncontrolled high blood pressure have spread throughout the world. The epidemics of obesity, diabetes, cardiovascular disease, depression, and disability are now global in scale, and both their incidence and prevalence are expected to increase as a result of the aging of the population and an exacerbation of health disparities. These risk factors and chronic diseases often occur in combination. For example, those with a poor diet and who are physically inactive are at elevated risk for diabetes, cardiovascular disease, and mood disorders.
Each disease is treated with multiple medications and psychotherapies. In the U.S., approximately 92% of older adults have at least one chronic condition and 62% have multiple conditions.27 Furthermore, 87% of those aged >65 years take prescription medications, and among those, the average number of such medications daily is four.28 This same population sees seven different health-care providers across four separate practices each year.29 Moreover, even though medical records are increasingly digitized, use of these to coordinate, let alone optimize, medical care appears to be many years away.29,30 Many question the sustainability of these current approaches to chronic disease, in particular with medical costs now exceeding $3 trillion/year in the U.S., with projections of >$5 trillion/year by 2023.31 If this is not sustainable, then other approaches must be found, including care provided by digital interventions.

The second trend is that the foundational principles are changing by which the determinants of health and health behaviors, and the design of health interventions to improve health outcomes, are understood. Notions of disease etiology and progression grounded in periodic assessment of biomarkers and physiological measures taken in medical office visits, or based on self-report of behaviors or environmental exposures, are in decline.32 New methods of understanding health and illness are emerging that are based on objective data on the genome, behaviors, social networks, psychological factors, social determinants, and the environment. For example, with respect to obesity, the recent discovery of the human microbiome33 and its potential relationship to obesity34 and obesity-related issues such as physical activity35 and energy balance36 increases the emphasis on the need for a systems biology approach to the problem37 that incorporates environmental influences on health into the equation.38 The aforementioned new digital health ecosystem allows data to be drawn in as needed from relevant areas and processed in real time through techniques like machine learning to yield predictions of health states and behavioral phenotypes.39 Integration of these markers collected from the context of people’s everyday lives with genomic and clinical databases is on the near horizon.40

The third trend is that health care is becoming more user-centered. This is changing the role of both doctors and patients as they learn, more than ever before, to collaborate, interpret, and act on shared sources of information that promote patient self-monitoring and self-management.41,42 These developments require new ways to view data privacy and health data ownership in which patients, or collectives of patients, become the owners of their own data and thus increase their say in how health care is provided and how health research is prioritized by whom and for whom.

### Health Behavior Research: New Data, Research Designs, and Methods

Technologic advances now illuminate what has been long theorized about behavior, that it is influenced at multiple levels—genetic, biological, social, environmental—and that these influences are reciprocal, dynamic, and temporally based.43,44 Thus, the complexity of understanding behavior strains current scientific methods and processes—something that is labeled “data poor.” A data-poor science requires researchers first to specify the questions, design the study to answer these questions, and then expend considerable time and resources to obtain study participants and collect data as per a prespecified study design. This approach provides considerable data control but is inefficient and often results in findings that are dated or obsolete, especially for the fast-paced technologic age.45 The enormous data collection effort is usually used to answer only the question and then the data are seldom shared and usually never used again. Moreover, it fails to produce a cumulative science in which the next iteration of the intervention builds on prior testing of the intervention.46 Further, this data-poor science places greater emphasis on “on average” insights. This is counter to the emphasis of the recent Precision Medicine Initiative47 that is focused on individual, contextual, and “lifestyle” factors that influence prevention and treatment.

Data-rich science, on the other hand, supports moving beyond single-level to multilevel models of analysis, and replace coarse and long timescale predictions with refined and multi-timescale predictions that match the complexity of the interactions among behaviors, their influences, and health outcomes. It also provides the ability to implement and rapidly evaluate models of behavior and interventions (the latter further described below in the section on agile science). An exemplar of this process is the field of meteorology that over its history48 has leveraged five technologies and processes to make this transition:

1. **New communication technologies.** In early meteorology, ground-based weather measurements were independent and isolated from each other, limiting it to descriptive surveillance and annual predictions for each location (e.g., Farmer’s Almanac). Continuous advances in communications, from the telegraph to the Internet, now allow meteorology to knit together a network of data stations and share data between them. This includes citizen science–driven approaches to data acquisition and from things like the increasingly ubiquitous weather stations in K–12 schools.

2. **Data standards.** As the sharing of meteorologic data increased, the need to standardize the data shared...
increased. By ensuring the data were comparable across these different measurement sites, more robust models could be generated from the data. This data sharing was particularly supported when all weather sensor systems were linked via an early version of the Internet (i.e., the AFOS computer system in 1979).

3. Multilevel data collection. Extending data collection to “multilevel,” initially by use of weather balloons and later by satellite imagery improved explanation and prediction by incorporating “higher-level” influences of the weather in the computational models.

4. Computational resources. With more data from more sources, there was an increasing need for computers that could rapidly manage the data, particularly within complex modeling frameworks such as dynamic systems. These models allowed for increasingly more accurate predictions especially as the feedback loop between prediction and outcome was greatly shortened.

5. Iteratively refining and optimizing multiple competing computational models. As computational capacity increased, a wider range of data modeling techniques emerged that further enabled more accurate and precise predictions. Iteratively testing multiple competing computational models and incorporating a human “in the loop” of the predictions to further refine them resulted in robust feedback loops for translating past observations into actionable knowledge such as daily weather forecasts.

For the behavioral sciences to transition to a multilevel, multi-timescale predictive science, it similarly needs to utilize these approaches to produce highly individualized knowledge about how to improve health behaviors. It could be argued that today’s current behavioral theories are akin to the Farmer’s Almanac as they are largely descriptive, past-oriented, and simplified to a few elements. These models for understanding behavior and behavior change provide largely “on average” insights without the level of specification and prediction that could occur in behavioral science if the approach to communication, data, and iterative evaluation of computationally complex, multilevel models now common in meteorology could be replicated. For example, it is well known that physical activity fluctuates in the short term (e.g., day-to-day) and long term (e.g., over a lifetime), with many individuals attempting but not sustaining changes in physical activity. Periodic behavioral surveys or measurements such as the Behavioral Risk Factor Surveillance Survey and National Health and Nutrition Examination Survey are not capable of tracking these changes, whereas the proliferation of new and more accurate wearable devices described above can.

Communication, Data, and Analysis
Behavioral science can be based in a data-rich research infrastructure that enables data sharing and standardization. Ecologically valid data about behaviors can be made available for use by researchers, patients, and individuals to ask questions and share insights into what supports and sustains healthful behavior change. As sources of these data grow, merge, and become increasingly temporally dense, the approach to research design and data analysis must change. Traditional research methods focus on manipulating an independent variable to isolate its effects on the dependent variable to test an a priori hypothesis. Nearly all of the other variables that might “confound” the impact of the independent variable on the dependent variable are relegated to the error term via random assignment to independent variable conditions. These new complex and extensive data sets allow for many more variables to be considered, assessments of their relative contribution on the behavior of interest, and the dynamic interplay of these variables over time. The accumulation of knowledge includes both discovery and confirmation, and thus the Bayesian style of statistics may become more useful than frequentist or null hypothesis statistical testing approaches. A Bayesian approach, by design, supports incremental model building via the use of “priors” knowledge. When a new data set is gathered, it can be compared to this to develop a “posterior” estimate, which can then be used as the next prior, and the cycle continues.

As the models evolve from descriptive to predictive and are iteratively improved, they support describing the association of variables over time and predicting how the magnitude and timing of the change of any one of these variables in the system affects all of the other variables. An example of such an approach to the unguided treatment of depression has been applied in the European FP7 ICT4DEPRESSION project, which is currently being evaluated in the E-COMPARED project (www.e-compared.eu) in which in the conduct of RCTs is combined with predictive modeling of both health economic costs and individual patient progress based on Markov, discrete event simulation, and Bayesian modeling techniques. By combining these approaches, this project can both predict what the outcome will be on average at the group level compared with non-intervention or treatment as usual as well as for which group of patients/users these interventions will be effective and what kind of action
could be undertaken beyond the individual intervention level.

Finally, Bayesian analyses can support facets of a problem that may not be as well represented in data such as subjective utility functions. This is critical for developing behavioral interventions for real-world use, as it supports a decision framework that can better balance the competing values of effectiveness at achieving a target goal while also accounting for issues of usability, safety, and cost.\(^{21,22}\) For example, in the physical activity scenario described above, incorporating subjective information is important if sustained improvements in physical activity are to be achieved as individuals experience transitions in life such as marriage, childbirth, and relocation to new environments.

**Agile Science and Iterative Evaluation**

The emergence of new technologies, new technology infrastructures, and new data sources have given rise to many intervention apps, wearables, and devices that target behaviors to support health, mental health, and wellness. Though great in numbers, the quality and evidence base for these are noticeably lacking.\(^2\) This is due, in part, to the fact that current scientific methods and practices are not capable of focusing on the necessary simultaneous and multiple iterative “trials” needed to define, refine, and optimize behavior change interventions. Instead of a phase-based model in which interventions are developed and then tested, multicomponent, multilevel, and data-rich interventions need to be continuously iterated, improved, and optimized. To accomplish this requires adapting methods from the engineering community in which development and evaluation occur in parallel, synergistically and iteratively until the solution has been optimized, a process called agile science.\(^{55}\) Agile science is intended to provide a framework for rapid iteration and improvement of systems before they are widely deployed.

Agile science focuses on identifying the most important assumptions currently made about a problem, goal, or solution, and then utilizing the most resource-efficient strategy possible to evaluate these assumptions to support decision making. The aim is to achieve rigor through a far greater emphasis on “trial and error” style science whereby many tests are run quickly and efficiently, with less emphasis placed on any one trial and more emphasis on a form of rapid replication. This iterative process incrementally builds on past successes while also identifying plausible but ultimately dead-end lines of inquiry. Agile science emphasizes the development of three “knowledge products:” modules, computational models, and personalization algorithms, which are incrementally and iteratively developed through an initial “sprint” phase, followed by an optimization phase, and finally an open source “release” phase.

Modules are the fundamental building blocks of this approach to behavioral interventions, as they represent mechanisms that support behavior change. For example, one module could be an adaptive goal-setting intervention that defines a daily “ambitious but doable” goal for an individual based on past behavior, psychosocial variables (e.g., stress, social interactions), and context (e.g., location). Computational models are used to predict how modules, individuals, and context might interact with novel users and contexts, and personalization algorithms translate the modules and computational models into dynamic decisions rules to support individuals in changing their behavior.\(^{55}\) Modules, when well validated, can be used to develop and test more-robust computational models and personalization algorithms and provide the necessary building blocks for the sort of personalized and perpetually adapting interventions described below.

The authors envision that multiple competing models can be evaluated simultaneously against one another on their ability to support improved prediction and regulation of a specific individual’s behaviors rather than simply “on average” insights, a concept called “idiographic generalization.”\(^{23}\) The iterative development of these knowledge products within a data-rich ecosystem promises to enable behavioral science to move toward predictive, multilevel, and multi-timescale models of behavior change that can drive precise behavioral interventions for each individual. To work well, this role will be informed by human input for support that is either beyond the capability of the system or not preferable to the end users.

Beyond predictive modeling techniques, other approaches will be needed to tackle different facets of this complex multidimensional problem of behavior change if it is to reflect the user-centered and collaborative care of the future. For example, machine learning may play an essential role in supporting pattern recognition within complex data sets. For example, it is becoming increasingly common for machine learning techniques to cull through very large data sets of multiple individuals to identify meaningful patterns of behavior.\(^{36,57}\)

Applying the agile science process in behavioral science is still in its formative stage, and many questions remain about its utility. However, as reflected in the other papers in this series, significant advances in behavioral science research are being made and the authors believe that the proposed methods and processes are compatible with this work. Moreover, the authors believe that this new approach to behavioral science is essential to the optimization and compatibility of behavioral
interventions within learning healthcare systems.29,58
Given appropriate standards of measurement and ontologies, and an increasingly powerful knowledge base, agile techniques can be used to iteratively improve system inputs and processes to achieve desired health outcomes for individuals and populations.

Implications for Intervention Development and Evaluation

Analyzing temporally dense data across levels not only improves explanation and prediction but also can serve as the basis for interventions that intervene at points in time considered optimal for behavior change.59 Historically, behavioral interventions addressed all individuals and contexts. Then came personalized and tailored interventions based on the baseline characteristics of the individual. Although tailored interventions initially met with disappointing results in both offline and online interventions, more recent work using web-based tailored interventions based on multiple theory-based moderators simultaneously have shown improved outcomes on health behavior.60

With the technologic revolution in data inputs, interventions can now be adapted initially and through the treatment course to changes in the individual and context.23 Early applications of such ecologic momentary interventions61 focused on the ability of mobile technologies to push out interventions throughout the day, sometimes with the advantage of recent data inputs or environmental cues, but often with only time as an adaptation variable (e.g., send dietary interventions around meal times).62 More recently, intensively adaptive interactions take into account “time-varying moderators” such as stress or activity to adapt the intervention63 and newly collected data increasingly drive the intervention. Just-in-time adaptive interventions are being tested that quickly sample and use data on an individual’s current state, situational context, and prior intervention experiences to deliver the most appropriate intervention at the optimal time.64 For example, a recent study showed the advantage of an adaptive system called “MyBehavior” that provided context-dependent suggestions for walking to improve physical activity (e.g., You are about to walk to this location, try this slightly longer route to fit in more activity).55 Just-in-time adaptive interventions go beyond intensively adaptive interventions: The intervention is adjusted daily or more often based on the real-time data being collected66 as they collect and analyze intensive longitudinal data to develop predictive algorithms that optimize the content and timing of the intervention and potentially pre-empt behaviors before they occur.

For many conditions, particularly those that are chronic and require sustained effort at behavior change such as obesity and mental health, current intervention technologies frequently fail to provide clinical benefit in real-world settings. This is primarily because people do not use them.67,68 Having a human presence, such as a provider, coach, or peers can enhance both use and efficacy.69–73 However, it is possible that the functions of human coaching, such as providing accountability and support,22 may be automated through virtual conversational agents.74,75

Although digital interventions have shown promise in pilot studies, current evaluation methods are undeveloped or perform poorly. Though methods are needed that protect users and other stakeholders and ensure efficacy, safety, and cost effectiveness,21 standard evaluation methods, such as RCTs that lock down interventions, are often ill suited for this new rapidly changing field.11 A number of solutions have been proposed, including methods that allow for iteration and learning during a trial,76 adaptive designs, regression discontinuity designs, A/B testing, open source platforms,77 or those that use post-marketing surveillance to monitor safety and effectiveness,78 and others discussed in this set of papers.22

Conclusions

This is a time of three major trends: increasing capabilities inherent in communication, computing, and data science; unsustainable growth in the complexity and cost of health care; and a movement to a more user-centered and collaborative approach to health promotion and health care. As outlined in this paper, the first and third trends can be leveraged to help address the second if public health is open to incorporating models of research and practice that are already being used in other data-intensive domains. The approach advocated in this paper is a radical departure from business as usual in behavioral science, and although well-founded concerns regarding such a departure need to be addressed, as this and other papers in this series outline, there are compelling reasons to increase efforts to explore the potential provided by these new technologies to transform science and improve public health.

This 2016 theme section of the American Journal of Preventive Medicine is supported by funding from the NIH Office of Behavioral and Social Sciences Research (OBSSR) to support the dissemination of research on digital health interventions, methods, and implications for preventive medicine.

This paper is one of the outputs of two workshops, one supported by the Medical Research Council (MRC)/National Institute for Health Research (NIHR) Methodology Research Program (PI Susan Michie), the OBSSR (William Riley, Director) and the Robert Wood Johnson Foundation (PI Kevin Patrick); and the other by the National Science Foundation (PI Donna Spruitj-Metz, proposal # 1539846).
David Mohr discloses that he has received honoraria from Otsuka Pharmaceuticals within the past 12 months. No other financial disclosures were reported by the authors of this paper.

References

29. IOM. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America - Institute of Medicine. 2012.
36. David Mohr discloses that he has received honoraria from Otsuka Pharmaceuticals within the past 12 months. No other financial disclosures were reported by the authors of this paper.

www.ajpmonline.org


