NODE – Review Article

An Algorithm for Digital Medicine Testing: A NODE.Health Perspective Intended to Help Emerging Technology Companies and Healthcare Systems Navigate the Trial and Testing Period prior to Full-Scale Adoption

Amy R. Sheon[a]   Brian Van Winkle[b]   Yauheni Solad[c]   Ashish Atreja[d]

[a]Case Western Reserve University School of Medicine, Rocky River, OH, USA; [b]Sibley Memorial at Johns Hopkins Medicine, Washington, DC, USA; [c]Yale New Haven Health, Stratford, CT, USA; [d]Mount Sinai Hospital, New York, NY, USA

Abstract
Over the last several years, there has been rapid growth of digital technologies attempting to transform healthcare. Unique features of digital medicine technology lead to both challenges and opportunities for testing and validation. Yet little guidance exists to help a health system decide whether to undertake a pilot test of new technology, move right to full-scale adoption or start somewhere in between. To navigate this complexity, this paper proposes an algorithm to help choose the best path toward validation and adoption. Special attention is paid to considering whether the needs of patients with limited digital skills, equipment (e.g., smartphones) and connectivity (e.g., data plans) have been considered in technology development and deployment. The algorithm reflects the collective experience of 20+ health systems and academic institutions that have established the Network of Digital Evidence for Health, NODE. Health, plus insights from existing clinical research taxonomies, syntheses or frameworks for assessing technology or for reporting clinical trials.

© 2018 The Author(s)
Published by S. Karger AG, Basel

A.R.S. is the lead author and B.W., Y.S., and A.A. are co-authors.

Brian Van Winkle, MBA, Executive Director of Innovation
Sibley Memorial at Johns Hopkins Medicine
5255 Loughboro Road NW
Washington DC 20016 (USA)
brianvanwinkle @ gmail.com

KARGER

Sheon et al.: Optimizing Digital Medicine Testing

Introduction

Over the last several years, there has been rapid growth of digital technologies aimed at transforming healthcare [1, 2]. Such innovations include patient-focused technologies (e.g., wearable activity tracking devices), other connected devices (glucometers, blood pressure cuffs, etc.) as well as the broader category of analytic applications used by health systems to monitor disease status, predict risk, make treatment decisions or monitor administrative performance. The emergence of nontraditional players in healthcare (e.g., Amazon and Google) [3, 4] is creating tremendous pressure on health systems to innovate. Yet little guidance exists to help a health system decide whether to undertake a pilot test of new technology, move right to full-scale adoption or somewhere in between [5].

Unique features of digital medicine technology lead to both challenges and opportunities for testing and validation. To navigate this complexity, this paper proposes an algorithm (Fig. 1) to help health systems determine the appropriate pathway to digital medicine validation or adoption. The algorithm was developed based on the collective experiences of 20+ health systems and academic institutions who have established the Network of Digital Evidence for Health, NODE.Health. In addition, evidence-based clinical research taxonomies, syntheses, or frameworks for assessing technology or for improving reporting of clinical trials identified from a literature review framed our work. Informed by Realist Synthesis methods [6, 7], the Immersion/Crystallization method [8] was used to understand how a product works by looking at the relationships among the product, the context, the user population, and outcomes. This approach is especially relevant to digital technology where complex interventions are continuously updated to reflect the user and context [9]. Concepts emerging from our literature review of contextual issues include classification of users [10–12], use of technology for collecting outcomes [13], and factors associated with scale-up [14, 15]. Concepts related to understanding outcomes were found in the literature related to methods of usability testing [16, 17], research design [18–22], product evaluation or efficacy testing [23–26], and reporting of research [24, 27–30]. To better understand the mechanism of action, we considered classification of technology [31–35], patient engagement or technology adoption [36–42], and reasons for nonuse of technology [43].

To make effective use of the algorithm, some key features of digital medicine products and their development ecosystem are discussed below. The user-task-context (eUTC) framework provides a lens for considering choices or plans for technology testing and adoption [44]. First, the institution should define its use case(s) for the product – know the target population (users), the required functions (tasks), and understand what is required for successful adoption within a specific setting. In assessing the fit of a technology under consideration, one should consider whether a product has been shown to work in the target user population. User dimensions especially important to health technology adoption include age, language, education, technology skills, and access to underlying necessary equipment and connectivity (e.g., smartphones or computers and broadband or mobile data). Similarly, one must ensure that a technology successfully used in one setting was used for the same tasks as are needed in the target situation. For example, reports may show that a program performed well with scheduling in a comparable setting but that no evidence was provided about its use for referrals, the task required by the target setting. Finally, numerous dimensions of the setting could affect the likelihood of success; consider, for example, the physical location, the urgency of the situation [44] and the social context (e.g., whether researchers or clinicians are involved, and their motivations for doing so) [45].

The potential for a misalignment of goals and incentives of institutional customers and technology developers and vendors is a dimension of the social context that is unique to
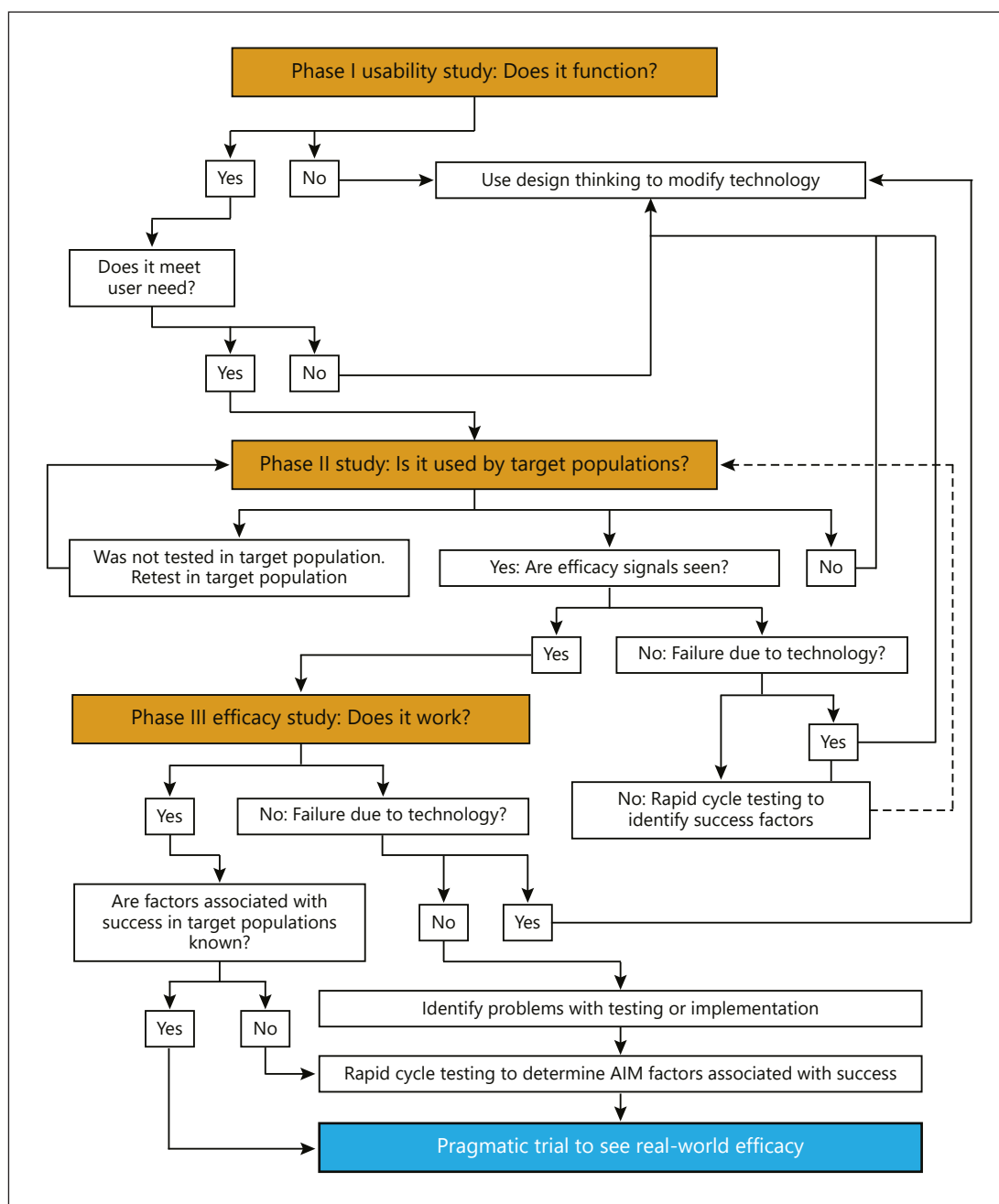
Sheon et al.: Optimizing Digital Medicine Testing

Fig. 1. The digital health testing algorithm.

digital health products. For example, a developer may encourage pilot studies as a prelude to making a lucrative sale but ultimately be unable to execute a deal because the pilot testers lack the authority or oversight to commit to an enterprise-wide purchase [46]. Or, health systems might want to test modifications of existing solutions to increase impact, but the needed updates may not fit with the entrepreneur's business plan. Finally, there must be adequate alignment between the institution and the product developer in terms of project goals, timeline, budget, capital requirements, and technical support.

KARGER

Regulatory Perspectives Associated with Digital Medicine Testing

The traditional clinical testing paradigm does not neatly apply to digital medicine because many digital solutions do not require regulatory oversight [47]. Mobile applications that pose a risk to patient safety [48], make medical claims, or are seeking medical grade certification [49] invoke Food and Drug Administration (FDA) oversight. For others, the agency reserves "enforcement discretion," meaning it retains the right to regulate low-risk applications [50] but allows such products to be sold without explicit approval. Senior US FDA officials acknowledge that the current regulatory framework "is not well suited for software-based technologies, including mobile apps…." [51]. The agency is gradually issuing guidance to address mobile health applications that fall in the gray area. For example, the FDA recently stated that clinician-facing clinical decision support will not require regulation, but that applications with patient-facing decision support tools may be regulated [52]. In light of the shifting landscape and gaps in regulatory oversight of digital health application, we review some unique features of digital health technologies and then offer an adaptation of the traditional clinical trial phases to provide a common language for discussing testing of digital medicine products.

Digital Medicine Product Characteristics

Digital health tools have some inherent advantages over traditional pharmaceutical and device solutions in that they may be commercially available at low cost, can undergo rapid iteration cycles, be rapidly disseminated through cell phones or the internet, can collect data on an ongoing rather than episodic basis, and can make data available in real time [53]. Digital tools can be highly adaptive for different populations, settings, and disease conditions. However, these very features create challenges for testing because digital tools can be used differently by different people, especially in the absence of training and standardization [53].

Digital Medicine Development Ecosystem

Much digital medicine technology is developed in the private sector where secrecy and speed trump the traditional norms of transparency and scientific rigor, especially in the absence of regulatory oversight. The ability to rapidly modify digital medicine software is essential for early-stage technology, but if algorithm updates are unbeknownst to evaluators, they can upend a study. On the other hand, in light of the average 7-year process to develop, test and publish results from a new technology, failure to update software could lead to obtaining results on an obsolete product [21, 53]. Digital products based on commercial off-the-shelf software are especial likely to undergo software changes versus digital solutions that are being built exclusively for a designated setting. Thus, it is important to understand where a technology is in its life cycle as well as to know whether the testers have a relationship with the developers whereby such background software changes would be disclosed or scheduled around the needs of the trial.

The type of company developing the digital technology can also affect testing dynamics. Start-up companies may have little research experience and limited capital to support operations at scale. Giants in the technology, consumer, transportation or communication sectors that are developing digital health tools may have the ability to conduct research at a massive scale [54, 55] (e.g., A/B testing by a social media company) but without the transparency typical of the established scientific method. Both types of companies may lack experience

Sheon et al.: Optimizing Digital Medicine Testing

Table 1. Special considerations for working with start-ups and nontraditional[a] firms

Topic	Questions to ask
Company and product maturity	How consistent is the proposed activity with the priorities of a company (ideally including founders, investors and other players who may influence the long-term relationship with the company)? Can product development survive key personnel transitions? How mature is the product? Can it remain stable for study duration? Does the company have the capital needed to provide adequate product and support for the duration of the trial and beyond? Does the company expect that the healthcare system will become a paying customer after the pilot phase? Who will pay for the product during and after a study? Will the company be able to support implementation if the testing progresses to adoption and scale-up? How many paying customers does the company currently have and have they undergone any types of studies already?
Fit with healthcare	Can the company provide the secure environment required to protect patient data and follow current regulations? Does the company understand the timeline, cost and cultural norms involved in working with health care institutions?
Research experience	Does the company understand requirements associated with undergoing an IRB review? Does the company understand the need to adhere to a protocol? Does the company understand requirements regarding human subject protection in research studies and health data privacy? Does the company understand issues with actual or appearance of conflict of interest in research?

[a] Firms with limited experience working with the healthcare sector. Internal divisions or departments of established organizations that work on novel approaches or technologies should be considered as start-ups.

working with the health sector and thus be unfamiliar with bifurcated clinical and administrative leadership.

In light of the many unique dimensions of digital health technology, Table 1 presents questions to consider when initiating or progressing with a phased trial process. These questions have been informed by the experience of NODE.Health in validating digital medicine technologies [47, 56–59], and by industry-sponsored surveys of digital health founders, health systems leaders, or technology developers [60, 61]; or other industry commentary [46, 62].

Testing Drugs and Devices versus Digital Solutions

Table 2 proposes a pathway for testing digital health products that addresses their unique characteristics and dynamics. The testing pathway used for regulated products is shown on the left with parallel digital testing phases shown on the right half of Table 2.

Preclinical studies, often performed in animals, are meant to explore the safety of a drug or devices prior to administration in humans [63]. The digital medicine analog is either absent or perhaps limited to determination of whether or not a product uses regulated technology such as cell phones that emit radiation.

Traditional and digital phase I studies may be the first studies of a drug or a technology in humans, are generally nonrandomized, and typically include a small number of healthy volunteers, or individuals with a defined condition. Whereas the purpose of traditional phase I studies is to test the agent's safety and find the highest dose that can be tolerated by humans

Sheon et al.: Optimizing Digital Medicine Testing

Table 2. Crosswalk of clinical and digital medicine testing phases

Drug and device trials			Digital medicine studies		
therapeutic and device testing phases			digital health (nonregulated) product testing phases		
phase	purpose	population	phase	purpose	population
Preclinical	Test drug safety and phar-macodynamics Required to test new drug in humans	Animals	Pre-marketing	Classify device accessories into levels of risk (e.g., class I or class II) Ensure safety of radiation-emitting devices	Animals or lab only
Phase I: in vivo	Safety and toxicity	Healthy or sick volunteers (n = very small)	Phase I: in silico	Identify likely use cases Rapidly determine whether and how a product is used Identify desired and missing features	Healthy or sick volunteers (n = small)
Phase II: therapeutic exploratory	Safety, pharmacokinetics, optimal dosing, frequency and rout of administration, and endpoints Generate initial indication of efficacy to determine sample size needed for phase III	Sick volunteers (n = small)	Phase II: feasibility	Develop/refine clinical work flow Assess usability Generate initial indication of efficacy to determine sample size needed for phase III	Volunteers reflecting Intended users (n = medium)
Phase III: pivotal or efficacy	Experimental study to demonstrate efficacy Estimate incidence of common adverse events	Diverse target popu-lation (n = medium)	Phase III: pivotal or efficacy	Experimental study to demonstrate clinical effectiveness and/or financial return on investment	Intended users, defined settings and context (n = medium/large)
Phase IV: therapeutic use or postmar-keting	Observational study to identify uncommon adverse events Evaluate cost or effec-tiveness in populations	Reports collected from universe of real-world users (n = census)	Implemen-tation	Observational or experimental design Uses implementation science including change management and rapid cycle testing to deploy a tested intervention at scale Assess effectiveness and unintended conse-quences in specific or real-world settings Assess integration requirements, interoper-ability, staffing and staff training require-ments for widespread adoption	Intended users, real-world setting and context (n = universe in a specific setting)
Pragmatic (effectiveness) trial	Understand effectiveness of drug or device in real-world settings	May use EHR or other "big data" sources Diverse settings and patient population	Pragmatic (effec-tiveness) trial	Understand effectiveness of technology in real-world settings	May use EHR or other big data sources Diverse settings and patient population May be conducted completely independently from healthcare setting

[63], the digital equivalent usually begins with a prototype product that may not be fully func-tioning. These early digital phase I studies increasingly use "design thinking" [16, 64] to identify relevant use cases and rapidly prototype solutions. Through focus group discussions, interviews, and surveys, developers find out what features people with the target condition want and conduct usability testing to see how initial designs work.

Traditional phase II studies look at how a drug works in the body of individuals with the disease or condition of interest. By using historical controls or randomization, investigators can get a preliminary indication of a drug's effectiveness so that endpoints and the sample size needed for an efficacy trial can be identified [63]. These techniques may require that study volunteers be homogenous, limiting understanding of product use among the full range of target product users. The digital analogue could be a feasibility study in which an appli-cation is tested with individuals or in settings reflecting the targeted end users. Because nonregulated digital products are by definition at low or no risk of harm, phase II studies of digital products may involve larger numbers of subjects than traditional trials.

Phase III trials of traditional and digital products are typically randomized studies to determine the efficacy of the drug or digital intervention. Ideally, volunteers reflect the full range of individuals with the condition, but, at least with therapeutic and device trials, the

Sheon et al.: Optimizing Digital Medicine Testing

study population is likely to be homogeneous with respect to variation in disease severity, comorbidities and age to control main sources of bias and maximize the ability to observe a significant difference between the treatment and control arms [65]. Digital studies may have fewer eligibility criteria resulting in studies that could be more broadly generalizable.

Because traditional efficacy trials are not likely large enough to detect rare adverse events, the FDA may require postmarketing (phase IV) studies to detect such events once a drug or device is in widespread use [65]. For low-risk applications, postmarketing surveillance studies may be unnecessary. Instead, we propose a class of studies that apply implementation science methods such as rapid cycle testing to generate new knowledge about technology dissemination beyond the efficacy testing environment and identification of unintended consequences in specific real-world settings. The trial may involve assessing requirements needed to integrate the novel product with existing technology, determine staffing requirements, training, and workflow changes. In this category, we also include strategies that allow for testing multiple questions in a short period of time, such as "n of one" and factorial studies [21]. Again, the Realist Synthesis perspective points to using these studies to best understand how and why a product does or does not work in specific settings [9].

Pragmatic trials are a next step common to both traditional and digital interventions; they seek to understand effectiveness in a real-world setting and with diverse populations [22, 66]. Pragmatic studies of digital technology could be undertaken completely outside of the context of healthcare settings such as with studies that recruit participants and deliver interventions through social media [67], that deliver interventions and collect data from biometric sensors and smartphone apps [68], or that market health-related services such as genetic testing directly to consumers [69]. The FDA's move toward the use of real-world evidence and real-world data suggests that digital health tools used by consumers may assume greater importance for testing drugs and devices [65] as well as nonregulated products.

Focus on External Validity

Many digital solutions can be widely adopted because they lack regulatory oversight and are commercially available to consumers. However, the specter of widespread utilization may magnify shortcomings of a testing process that may not have been conducted adequately in diverse settings and/or with desired groups of end users. The RE-AIM framework has been used since the late 1990s to evaluate the extent to which an intervention reaches, and is representative of, the affected population [70]. More recently, the framework has been used to plan and implement research studies [71] and to assess adoption of community-based healthcare initiatives [72]. Therefore, application of the RE-AIM framework to digital medicine testing can be especially valuable for those seeking to understand the likelihood that a specific digital health solution will work in a specific setting. Table 3 summarizes the original RE-AIM domains and then applies them to illuminate reliability and validity issues that could arise in testing digital medicine technologies.

The RE-AIM Dimensions and Key Questions

"Reach" concerns the extent to which populations involved in the testing reflect populations intended to use the product. "Effectiveness" concerns whether the product was effective overall or in subgroups typically defined by health status or age. Based on standards that have been articulated to improve the reporting of web-based and mobile health interventions [28] and reporting of health equity [29], we extend the traditional Reach and Effectiveness dimensions to highlight specific populations in whom digital medicine technologies might exhibit

Sheon et al.: Optimizing Digital Medicine Testing

Table 3. RE-AIM dimension and key questions for testing decisions

RE-AIM dimension and efficacy/effectiveness criteria	Original RE-AIM conception[a]	Digital health application
Reach	Percent and risk characteristics of individuals who receive or are touched by an intervention compared with an overall population How do participants differ from nonparticipants?	Who participated in testing a product versus who is the product intended for? Important population characteristics to consider include: Health condition (e.g., chronic vs. acute) Demographics (age, race, gender, education) Language and literacy (English, other, health literacy) Technology access (smartphone, computer, mobile data, fixed broadband, wi-fi) Technology skills
Efficacy	Assess positive and negative outcomes; include behavioral, quality of life and satisfaction in addition to health outcomes	How effective was the product overall? Were the intended benefits seen across various subgroups as defined above?
Adoption	The proportion and representativeness of settings that adopt a program. Reasons for nonadoption should also be assessed	In what setting (e.g., ambulatory care, in-patient, emergency department) was the product tested? How was it received by institutional users? Who (e.g., clinical, marketing, IT staff) needs to be involved in decisions to adopt the new technology? How much will it cost the institution and who will pay?
Implementation	The extent to which a program is delivered as intended. Outcomes should be assessed in terms of organizational delivery and individual use	How was the intervention delivered in trials? How was it received by end users? How long will it take to implement? How important is fidelity to the conditions seen in prior tests versus real-world use? What staff and training are required? How much will it cost end users?
Maintenance	The extent to which organizations routinize implementation of the program, and to which individuals maintain their use of the program or continue to experience benefits from the program	Is the intervention one-time, ad hoc or for ongoing use? Are effects expected to grow, plateau or diminish? How long are the results sustained? Are boosters and updates needed? Is staff onboarding and retraining needed?

[a] Source: Adapted from Glasgow et al. [70].

heterogeneity of effect due to "social disadvantage." Traditional social determinants of health (education, race, income) plus age are highly associated with access to and use of the internet and smartphones. Digital skills are, in turn, highly associated with having and using computers and smartphones, as such skills develop from use [73], and usage grows with access to equipment and connectivity [74]. Digital skills, computers or smartphone, and mobile data or fixed broadband are essential for almost all consumer-facing health technology such as remote monitors and activity tracking devices. Given the covariance of digital skills and access with age, education, and income [75–80], and the likelihood that those with low digital skills may resist adoption of a digital technology or may have challenges using it [56, 81–84], it is especially important to determine whether such individuals were included in prior testing of a product and whether prior studies examined heterogeneity in response by such categories.

For products that have proven effectiveness, "Adoption" concerns the incorporation of the intervention into a given context after pilot testing. Stellefson's application of the RE-AIM model to assessing the use of Web 2.0 interventions for chronic disease self-management in

Table 4. Application of RE-AIM framework to digital medicine testing by phase

Digital medicine testing phase	RE-AIM dimensions	Questions to ask existing evidence base	When to use
Phase I: use case identification	Reach	In which populations was technology tested? How similar were testers to current target population?	Developing a new product Using an existing product for a new indication or with a different population
Phase II: feasibility	Reach Effectiveness Implementation	In which setting was technology tested? How similar was study setting to target setting? How usable was the technology to the target population? To what extent were invited users interested in participating? What sort of support was needed from administrators, clinical staff and technology groups for the product to be tested?	Earlier tests were in populations or settings dissimilar to current target
Phase III: pivotal or efficacy	Reach Effectiveness Adoption Implementation	Was technology efficacious overall? Did study look for heterogeneity of results? Was technology effective in subgroups of interest? Was adherence adequate? Was there heterogeneity of adherence? Was testing of sufficient duration to know what will be needed for implementation, if efficacious? Are there study conditions (such as staff or patient training or executive championship) needed for the technology to be effective in real-world adoption? Does the study document financial outcomes that confirm a business case for product adoption?	Efficacy has not been proven in the population of interest ROI needs to be demonstrated for enterprise-wide adoption to be considered A target setting lacks characteristics that were necessary for efficacy in prior studies
Implementation	Reach Effectiveness Adoption Implementation Maintenance	How well was technology accepted by the institution? What factors were associated with successful and failed implementation? Did real-world efficacy match that seen in trials? Was adoption and efficacy consistent across settings and populations? Did usage improve and expand over time? Did technology adapt as needed over time?	There are no successful real-world adoption examples Target institution lacks success scaling up new technologies Target institution has settings and populations that are heterogeneous in factors connected with successful technology adoption
Pragmatic	Implementation Adoption Maintenance	What level of support is needed from the company to support implementation? How does the technology interact with existing systems? What support is needed to sustain and use and keep the technology updated?	Factors needed for successful implementation are known from other settings and replicable in the target setting Factors associated with successful adoption of other technologies in this setting are known and replicable

older adults is especially instructive for assessing how a healthcare institution moves to Adoption after a trial [85]. To what extent does the current setting replicate the context of previous trials? If a product was tested in an in-patient setting but is being considered for use in an ambulatory setting, for example, it is vital to understand who would need to be involved in a decision to move to adoption after efficacy testing. Do current staff have the skills and authority needed to integrate new technology? Where will funds for the technology adoption project come from?

"Implementation" refers to the types of support needed for the intervention to be administered and how well it was accepted by end users. "Maintenance" addresses the sustainability of the intervention in a setting and the duration of effect in the individual user.

We next apply the RE-AIM framework to help health systems assess evidence gaps and identify the most appropriate testing to use in a specific situation. For a product at an early stage of development, as shown in Table 4, questions of "Reach" are most relevant. Since early testing focuses on how well a device or application functions, there may be limited information available about how well the device functions among individuals or settings similar to those under consideration. For a technology that has completed an efficacy trial, it is important to understand the extent to which the test results would apply in different populations and settings. It is also crucial to understand features of the study such as whether careful instruction and close

monitoring were necessary for the technology to function optimally. For a technology with ample evidence of efficacy in relevant populations and settings, those considering moving to widespread adoption must understand institutional factors associated with successful deployment such as engagement of marketing and IT departments to integrate data systems.

The Digital Health Testing Algorithm

The digital health testing algorithm (Fig. 1) presents a series of questions to help a health system determine what sort of testing, if any, may be needed when considering adoption of a new technology.

Example 1: Care Coordination, Application, Validation

A healthcare system was approached by a start-up company that had designed a new care coordination application that uses artificial intelligence to predict readmission risk for a particular condition. The tool simplifies the clinical workflow by pulling multiple data sources together into a single dashboard. For this first attempt at use in a clinical setting, an ambulatory care practice was asked to check the electronic medical record to validate that the product correctly aligned data from multiple systems. With that validation step complete and successful, the start-up company would like to embark on a small prospective phase II study to validate the accuracy of the prediction model. Is that the right next step?

The initial study was a phase I study that showed that the data mapping process worked correctly using data pertinent to an ambulatory care setting. Further phase I testing should be done to ensure that the data mapping process works more broadly, such as with in-patient data. Then, to validate the algorithm, the company could apply the algorithm to historic data and see how well it predicts known outcomes. This would still be part of phase I, simply ensuring that the product works as intended. Once the key functions are confirmed to function, then additional qualitative studies should be undertaken to ensure that the product meets the needs of users. That process should lead to additional customization, based on real-time healthcare system feedback. Once that process is complete, they can move to phase II testing. For the first actual use in the intended clinical environment, the health system leadership might identify a local clinical champion who would be thoughtful and astute about how to integrate the product into the clinical workflow. Health system Information Technology staff should be closely involved to ensure a smooth flow of data for the initial deployment. From this stage, utilization data should be examined closely, and staff should be debriefed to determine whether it was used. Some initial signals that the product produces the desired outcome should be seen before proceeding to a phase III efficacy trial.

Example 2: Patient-Facing Mobile Application to Improve Chronic Disease Tracking and Management

An independent medical practice is approached by a start-up company that has developed a mobile application to simplify management of a particular condition. As no clinical advice is involved, the application is not regulated by the FDA. The product has been used for 1 year in three other small- to medium-size practices. Feedback from the practices and patients was used to tweak the product. A pilot study with the improved product showed that satisfaction and utilization were very high among patients who were able to download the app onto their phones. Hospitalizations were lower and disease control better among patients who used the application compared with nonusers, but the results did not reach statistical significance. The

Sheon et al.: Optimizing Digital Medicine Testing

start-up company has a Spanish language version of the product that was developed with input from native Spanish-speaking employees at the start-up company. They would like to conduct a phase III efficacy trial with a randomized design, using both English and Spanish language versions of the tool. Is that the right next step?

This product has completed phase II testing that showed efficacy signals among those who used the tool heavily. Before proceeding to an efficacy test in a comparable population, one must look carefully to see how patients who used the product differed from those who used it only a bit or not at all. If users are healthier and better educated, they may have been better able to manage their conditions even without the application. Qualitative assessments from staff and patients should be undertaken to ascertain why others did not use the product. If new functions are needed, they should be developed and tested in phase I studies. If operational barriers can be identified and addressed, rapid cycle testing could be undertaken, and then a repeat phase II study with the population that initially failed to use the product. In considering initiation of a trial in the new clinic, phase I studies must be undertaken to ensure that the product meets the needs of very different populations – including lower income, non-English-speaking, and patients with only rudimentary smartphones and limited skill in using apps. Then, a phase II study should be initiated to see whether the target population will use the app and whether efficacy signals are seen.

Example 3: Scale-Up of Disease Management System

A mid-size faith-based community healthcare system serving a largely at-risk population has implemented a disease-specific digital chronic care management and patient engagement program at two of its larger facilities. The deployment went well, and data strongly show that the application has positive outcomes. The system would like to roll the product out at all 20 of its facilities, using a pragmatic design to capture data on the efficacy of the product in real-world use. Should they move straight to a pragmatic trial?

In this case, 4 of the 20 additional clinics have newly joined this health system; various administrative systems are still being transitioned. If the health system has an excellent understanding of the contextual factors associated with successful implementation, then they may proceed with a pragmatic trial. For example, they may see that significant additional staffing is required to manage patient calls and messages that come through the app. They might then incorporate a call center to efficiently manage patient interactions at scale. However, if such factors are unknown, they should undertake rapid cycle testing, perhaps deploying the system in just one of the new clinics, and using qualitative research methods to understand factors associated with successful adoption. The focus of the rapid cycle tests and the pragmatic trial is on testing methods of scaling the product, rather than on testing the product itself.

Conclusion

Digital solutions developed by technology companies and start-ups pose unique challenges for healthcare settings due to potential for lack of alignment of goals between a health system and a technology company, the potential for unmeasured heterogeneity of effectiveness, and the need to understand institutional factors that may be crucial for successful adoption. Taxonomies and frameworks from public health and from efforts to improve the quality of clinical research publications lead to a set of questions that can be used to assess existing data and chose a testing pathway designed to ensure that products will be effective with the target populations and in target settings.

Sheon et al.: Optimizing Digital Medicine Testing

Some real-world examples [86, 87]87 illustrate the need for attention to the fit of a technology to the needed tasks, the user population, and the context of adoption.

Thies et al. [86] describe a "failed effectiveness trial" of a commercially available mHealth app designed to improve clinical outcomes for adult patients with uncontrolled diabetes and/or hypertension. Although tested very successfully in "the world's leading diabetes research and clinical care organization" and at one of the nation's leading teaching and research hospitals [88], use in a Federally Qualified Health Center failed for reasons clearly related to the "context" and the "users." Patients reported a lack of interest in the app as "getting their chronic condition in better control did not rise to the level of urgency" for patients that it held with the clinic staff. Lack of access to and comfort with technology "may have reduced the usability of the app for patients who enrolled and dissuaded patients who were not interested." Contextual factors included difficulty downloading the app due to limited internet access in the clinic, lack of clinic staff time to explain app use to patients, and lack of integration into clinic workflow since the app was not connected with the EHR [86].

In a pragmatic trial, 394 primary care patients living in West Philadelphia were called 2 days before their appointments with offers of Lyft rides. There were no differences in missed visits or in 7-day ED visits for those offered rides compared with those whose appointments were on different days where rides were not offered. Investigators had only been able to reach 73% of patients on the phone to offer rides, of whom only 36.1% were interested in receiving the offered Lyft ride. Hundreds of health systems have now deployed ride sharing services, but evidence of the impact has not been reported [87]. A considerable body of qualitative research may be needed to determine the best ways to deploy a resource that makes great sense intuitively but may not be received by users as envisioned.

Although far from complete, the proposed algorithm attempts to provide some structure and common terminology to help health systems efficiently and effectively test and adopt digital health solutions. In addition, a framework analogous to clinical trials.gov is needed to track and learn from the evaluation of digital health solutions that do not require FDA oversight.

Acknowledgement

Disclosure Statement

Brian Van Winkle is Executive Director of NODE.Health. Amy R. Sheon and Yauheni Solad are volunteer members of the NODE.Health Board of Directors. Ashish Atreja is the founder of NODE.Health.

Funding Sources

KARGER

roadmap for Medical Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. NODE.Health receives industry funding to evaluate the clinical effectiveness of digital solutions. NODE.Health is also developing a registry to track the reporting of digital health solutions to encourage greater transparency and better knowledge.

References

1 Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. mHealth. 2017;3:38.

2 Topol E. The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care. First Trade Paper Edition, Revised and Expanded edition. New York: Basic Books; 2013.

3 EHRIntelligence. How Google, Microsoft, Apple are impacting EHR use in healthcare. EHRIntelligence [Internet]. 2018 [cited 2018 Jun 19]. Available from: https://ehrintelligence.com/news/how-google-microsoft-apple-are-impacting-ehr-use-in-healthcare.

4 Report S. AI voice assistants have officially arrived in healthcare. Healthc. IT News [Internet]. 2018 [cited 2018 Feb 9]. Available from: http://www.healthcareitnews.com/news/special-report-ai-voice-assistants-have-officially-arrived-healthcare.

5 The Lancet. Is digital medicine different? Lancet. 2018 Jul 14;392(10142):95.

6 Pawson R. Evidence-based Policy: A Realist Perspective. London: Sage Publications; 2006. https://doi.org/10.4135/9781849209120.

7 Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review – a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy. 2005 Jul;10(suppl 1):21–34.

8 Borkan JM. Immersion/crystallization. In: Crabtree BF, Miller WL, editors. Doing Qualitative Research. 2nd ed. Newbury Park: Sage; 1999.

9 Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. How outcomes are achieved through patient portals: a realist review. J Am Med Inform Assoc. 2014 Jul-Aug;21(4):751–7.

10 Parasuraman A, Colby CL. An updated and streamlined Technology Readiness Index: TRI 2.0. J Serv Res. 2015;18(1):59–74.

11 Najaftorkaman M, Ghapanchi AH, Talaei-Khoei A, Ray P. A taxonomy of antecedents to user adoption years of research. J Assoc Inf Sci Technol. 2015;66(3):576–98.

12 Kayser L, Kushniruk A, Osborne RH, Norgaard O, Turner P. Enhancing the effectiveness of consumer-focused health information technology systems through eHealth literacy: a framework for understanding users' needs. JMIR Hum Factors. 2015;2(1):e9.

13 Perry B, Herrington W, Goldsack JC, Grandinetti CA, Vasisht KP, Landray MJ, et al. Use of mobile devices to measure outcomes in clinical research, 2010–2016: a systematic literature review. Digit Biomark. 2018 Jan 31;2(1):11–30.

14 Van Dyk L. A review of telehealth service implementation frameworks. Int J Environ Res Public Health. 2014;11(2):1279–98.

15 Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res. 2018;20(5):e10235.

16 Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res. 2016;18(12):e317.

17 Heffernan KJ, Chang S, Maclean ST, Callegari ET, Garland SM, Reavley NJ, et al. Guidelines and recommendations for developing interactive eHealth apps for complex messaging in health promotion. JMIR Mhealth Uhealth. 2016 Feb;4(1):e14.

18 Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth assessment: conceptualization of a global framework. JMIR MHealth UHealth. 2017;5(5):e60.

19 Mason VC. Innovation for medically vulnerable populations, Part 3: experiment design [Internet]. Tincture. 2016 [cited 2018 Feb 27]. Available from: https://tincture.io/innovation-for-medically-vulnerable-populations-part-3-experiment-design-47c8c7b00919#.r6wt2gwlb.

20 Malikova MA. Optimization of protocol design: a path to efficient, lower cost clinical trial execution. Future Sci OA. 2016 Jan 12;2(1):FSO89.

21 Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res. 2014 Feb 19;16(2):e36.

22 Porzsolt F, Rocha NG, Toledo-Arruda AC, Thomaz TG, Moraes C, Bessa-Guerra TR, et al. Efficacy and effectiveness trials have different goals, use different tools, and generate different messages. Pragmatic Obs Res. 2015 Nov 4;6:47–54.

23 Philpott D, Guergachi A, Keshavjee K. Design and validation of a platform to evaluate mHealth apps. Stud Health Technol Inform. 2017;235:3–7.

.·.Digital
Biomarkers

Digit Biomark 2018;2:139–154	
DOI: 10.1159/000494365	© 2018 The Author(s). Published by S. Karger AG, Basel www.karger.com/dib

Sheon et al.: Optimizing Digital Medicine Testing

152

24 van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res. 2011 Dec 5; 13(4):e111.

25 Monitoring and Evaluating Digital Health Interventions. A practical guide to conducting research and assessment [Internet] [cited 2018 Feb 5]. Available from: http://www.who.int/reproductivehealth/publications/mhealth/digital-health-interventions/en/.

26 Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res. 2012 Nov;14(6):e161.

27 Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ. 2016;352:i1174.

28 Eysenbach G; CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res. 2011;13(4):e126.

29 Welch VA, Norheim OF, Jull J, Cookson R, Sommerfelt H, Tugwell P, et al. CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials. BMJ. 2017;359:j5085.

30 Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016;355:i5239.

31 Hors-Fraile S, Rivera-Romero O, Schneider F, Fernandez-Luque L, Luna-Perejon F, Civit-Balcells A, et al. Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: A scoping review. Int J Med Inf. 2018;114:143–55.

32 Basher KM, Nieto-Hipolito JI, Leon MD, Vazquez-Briseno M, López J DS, Mariscal RB. Major existing classification matrices and future directions for Internet of Things. Adv Internet Things. 2017;07(04):112.

33 WHO. Classification of digital health interventions v1.0 [Internet]. WHO [cited 2018 Apr 5]. Available from: http://www.who.int/reproductivehealth/publications/mhealth/classification-digital-health-interventions/en/.

34 Donnelly JE, Jacobsen DJ, Whatley JE, Hill JO, Swift LL, Cherrington A, et al. Nutrition and physical activity program to attenuate obesity and promote physical and metabolic fitness in elementary school children. Obes Res. 1996 May;4(3):229–43.

35 Wang A, An N, Lu X, Chen H, Li C, Levkoff S. A classification scheme for analyzing mobile apps used to prevent and manage disease in late life. JMIR Mhealth Uhealth. 2014 Feb;2(1):e6.

36 Singh K, Drouin K, Newmark LP, Rozenblum R, Lee J, et al. Developing a Framework for Evaluating the Patient Engagement, Quality, and Safety of Mobile Health Applications [Internet]. 2016 [cited 2018 Mar 27].Available from: http://www.commonwealthfund.org/publications/issue-briefs/2016/feb/evaluating-mobile-health-apps.

37 Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol. 2016;6:2013.

38 Abelson J, Wagner F, DeJean D, Boesveld S, Gauvin FP, Bean S, et al. Public and patient involvement in health technology assessment: a framework for action. Int J Technol Assess Health Care. 2016;32(4):256–64.

39 Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc. 2009;16(4):550–60.

40 Kruse CS, DeShazo J, Kim F, Fulton L. Factors associated with adoption of health information technology: a conceptual model based on a systematic review. JMIR Med Inform. 2014;2(1):e9.

41 Karnoe A, Furstrand D, Christensen KB, Norgaard O, Kayser L. Assessing competencies needed to engage with digital health services: development of the eHealth literacy assessment toolkit. J Med Internet Res. 2018; 20(5):e178.

42 Sawesi S, Rashrash M, Phalakornkule K, Carpenter SJ, Jones FJ. The impact of information technology on patient engagement and health behavior change: a systematic review of the literature. JMIR Med Inform. 2016; 4(1):e1.

43 Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res. 2017;19(11):e367.

44 Kushniruk A, Turner P. A framework for user involvement and context in the design and development of safe e-Health systems. Stud Health Technol Inform. 2012;180:353–7.

45 Trivedi MC, Khanum MA. Role of context in usability evaluations: A review. Adv Comput Int J. 2012;3(2): 69–78.

46 Baum S. How can digital health startups steer clear of pilot study pitfalls? [Internet]. Medcity News. 2018 [cited 2018 Aug 17]. Available from: https://medcitynews.com/2018/03/how-can-digital-health-startups-steer-clear-of-pilot-study-pitfalls/?rf=1.

47 Makhni S, Atreja A, Sheon A, Van Winkle B, Sharp J, Carpenter N. The broken health information technology innovation pipeline: a perspective from the NODE Health Consortium. Digit Biomark. 2017;1(1):64–72.

48 Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff [Internet]. U.S. Food and Drug Administration; 2015 [cited 2016 Mar 29]. Available from: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf.

49 Sartor F, Papini G, Cox LG, Cleland J. Methodological shortcomings of wrist-worn heart rate monitors validations. J Med Internet Res. 2018;20(7):e10108.

KARGER

Sheon et al.: Optimizing Digital Medicine Testing

50 Larson RS. A path to better-quality mHealth apps. JMIR MHealth UHealth. 2018;6(7):e10414.

51 Shuren J, Patel B, Gottlieb S. FDA regulation of mobile medical apps. JAMA. 2018;320(4):337–8.

52 Clinical and patient decision support software: draft guidance for industry and Food and Drug Administration staff [Internet]. U.S. Food and Drug Administration; 2017 [cited 2018 Aug 23]. Available from: https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM587819.pdf.

53 Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med. 2013 Aug;45(2):228–36.

54 Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. Proc Natl Acad Sci USA. 2014;111(24):8788–90.

55 Cobb NK, Graham AL. Health behavior interventions in the age of Facebook. Am J Prev Med. 2012;43(5):571–2.

56 Sheon AR, Bolen SD, Callahan B, Shick S, Perzynski AT. Addressing disparities in diabetes management through novel approaches to encourage technology adoption and use. JMIR Diabetes. 2017;2(2):e16.

57 Winkle BV, Carpenter N, Moscucci M. Why aren't our digital solutions working for everyone? AMA J Ethics. 2017;19(11):1116–24.

58 Atreja A, Khan S, Otobo E, Rogers J, Ullman T, Grinspan A, et al. Impact of real world home-based remote monitoring on quality of care and quality of life in IBD patients: interim results of pragmatic randomized trial. Gastroenterology. 2017;152(5):S600–S601.

59 Atreja A, Daga N, Patel NP, Otobo E, Rogers J, Patel K, et al. Sa1426 Validating an automated system to identify patients with high risk lesions requiring surveillance colonoscopies: implications for accountable care and colonoscopy bundle payment. Gastrointest Endosc. 2015;81(5):AB210.

60 Evans B, Shiao S. Streamlining enterprise sales in digital health [Internet]. Rock Health; 2018 [cited 2018 Aug 14]. Available from: https://rockhealth.com/reports/streamlining-enterprise-sales-in-digital-health/.

61 Insights on Digital Health Technology Survey 2016: How digital health devices and data impact clinical trials [Internet]. Validic; 2016 [cited 2018 Aug 19]. Available from: https://hitconsultant.net/2016/09/19/validic-digital-health-devices-report/.

62 Surve S. How to survive IRBs and have successful digital health pilots [Internet]. Medcity News. 2017 [cited 2018 Aug 19]. Available from: https://medcitynews.com/2017/11/survive-irbs-successful-digital-health-pilots/?rf=1.

63 Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: a narrative review. Postgrad Med. 2011;123(5):194–204.

64 Dorst K. The core of "design thinking" and its application. Des Stud. 2011;32(6):521–32.

65 Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. JAMA. 2018 Sep 4;320(9):867–68.

66 Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63.

67 Napolitano MA, Whiteley JA, Mavredes MN, Faro J, DiPietro L, Hayman LL, et al. Using social media to deliver weight loss programming to young adults: Design and rationale for the Healthy Body Healthy U (HBHU) trial. Contemp Clin Trials. 2017;60:1–13.

68 Chen T, Zhang X, Jiang H, Asaeikheybari G, Goel N, Hooper MW, et al. Are you smoking? Automatic alert system helping people keep away from cigarettes. Smart Health [Internet]. 2018 [cited 2018 Aug 6]. Available from: http://www.sciencedirect.com/science/article/pii/S2352648318300436.

69 Tung JY, Shaw RJ, Hagenkord JM, Hackmann M, Muller M, Beachy SH, et al. Accelerating precision health by applying the lessons learned from direct-to-consumer genomics to digital health technologies. NAM Perspect [Internet]. 2018 [cited 2018 Sep 7]. Available from: https://nam.edu/accelerating-precision-health-by-applying-the-lessons-learned-from-direct-to-consumer-genomics-to-digital-health-technologies/.

70 Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health. 1999;89(9):1322–27.

71 Carlfjord S, Andersson A, Bendtsen P, Nilsen P, Lindberg M. Applying the RE-AIM framework to evaluate two implementation strategies used to introduce a tool for lifestyle intervention in Swedish primary health care. Health Promot Int. 2012;27(2):167–76.

72 Glasgow RE, Estabrooks PE. Pragmatic applications of RE-AIM for health care initiatives in community and clinical settings. Prev Chronic Dis. 2018 Jan 4;15:E02.

73 McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Entertainment or health? Exploring the internet usage patterns of the urban poor: a secondary analysis of a randomized controlled trial. J Med Internet Res. 2016 Mar 3;18(3):e46.

74 Schartman-Cycyk S, Meissier K. Bridging the gap: what affordable, uncapped internet means for digital inclusion [Internet]. Mobile Beacon; 2017 [cited 2017 May 20]. Available from: http://www.mobilebeacon.org/wp-content/uploads/2017/05/MB_ResearchPaper_FINAL_WEB.pdf.

75 Anderson M, Perrin A, Jiang J. 11% of Americans don't use the internet. Who are they? [Internet]. Pew Research Center. 2018 [cited 2018 Apr 13]. Available from: http://www.pewresearch.org/fact-tank/2018/03/05/some-americans-dont-use-the-internet-who-are-they/.

76 Internet/Broadband Fact Sheet [Internet]. Pew Research Center; 2017 [cited 2017 Nov 7]. Available from: http://www.pewinternet.org/fact-sheet/internet-broadband/.

77 Anderson M, Perrin A. Technology adoption climbs among older adults [Internet]. 2017 [cited 2017 May 20]. Available from: http://www.pewinternet.org/2017/05/17/technology-use-among-seniors/.

KARGER

Sheon et al.: Optimizing Digital Medicine Testing

78 Anderson M. Digital divide persists even as lower-income Americans make gains in tech adoption [Internet]. Washington (DC): Pew Research Center; 2017. Available from: http://www.pewresearch.org/fact-tank/2017/03/22/digital-divide-persists-even-as-lower-income-americans-make-gains-in-tech-adoption/#.

79 Horrigan JB. Digital readiness gaps [Internet]. Pew Research Center; 2016 [cited 2017 Feb 14]. Available from: http://www.pewinternet.org/2016/09/20/digital-readiness-gaps/.

80 Anderson M, Horrigan JB. Smartphones may not bridge digital divide for all [Internet]. Factank News Numbers 2016 [cited 2016 Nov 8]. Available from: http://www.pewresearch.org/fact-tank/2016/10/03/smart-phones-help-those-without-broadband-get-online-but-dont-necessarily-bridge-the-digital-divide/.

81 Perzynski AT, Roach MJ, Shick S, Callahan B, Gunzler D, Cebul R, et al. Patient portals and broadband internet inequality. J Am Med Inform Assoc. 2017 Sep 1;24(5):927–32.

82 Ackerman SL, Sarkar U, Tieu L, Handley MA, Schillinger D, Hahn K, et al. Meaningful use in the safety net: a rapid ethnography of patient portal implementation at five community health centers in California. J Am Med Inform Assoc. 2017 Sep 1;24(5):903–12.

83 Tieu L, Schillinger D, Sarkar U, Hoskote M, Hahn KJ, Ratanawongsa N, et al. Online patient websites for electronic health record access among vulnerable populations: portals to nowhere? J Am Med Inform Assoc. 2017 Apr;24 e1:e47–54.

84 Lyles C, Schillinger D, Sarkar U. Connecting the dots: health information technology expansion and health disparities. PLoS Med. 2015;12(7):e1001852.

85 Stellefson M, Chaney B, Barry AE, Chavarria E, Tennant B, Walsh-Childers K, et al. Web 2.0 chronic disease self-management for older adults: a systematic review. J Med Internet Res. 2013;15(2):e35.

86 Thies K, Anderson D, Cramer B. Lack of Adoption of a mobile app to support patient self-management of diabetes and hypertension in a federally qualified health center: interview analysis of staff and patients in a failed randomized trial. JMIR Hum Factors. 2017;4(4):e24.

87 Chaiyachati KH, Hubbard RA, Yeager A, Mugo B, Lopez S, Asch E, et al. Association of rideshare-based transportation services and missed primary care appointments: a clinical trial. JAMA Intern Med. 2018;178(3):383–9.

88 Moore JO. Technology-supported apprenticeship in the management of chronic disease [Internet]. 2013 [cited 2018 Aug 25]. Available from: http://hdl.handle.net/1721.1/91853.

KARGER